THE JOURNAL
OF
CHEMICAL PHYSICS

VOLUME 1 DECEMBER, 1933 NUMBER 12

The Derivation of Equations for Regular Solutions

J. H. HILDEBRAND AND S. E. WOOD, *Department of Chemistry, University of California*

(Received September 27, 1933)

The probability function of Menke has been used as a basis for a statistical treatment of the intermolecular potentials in solutions of symmetrical molecules for which the entropy of mixing is the same as for an ideal solution of the same composition. The resulting expressions enable calculations to be made of the departure of iodine solutions from the ideal solution laws which agree satisfactorily with the observed departures.

THE senior author has called attention[1] to the existence of families of solubility curves for a single solute in a number of solvents where both components are comparatively nonpolar and the chemical effects of solvation, association and compound formation are absent. Such solutions have been called regular, and it has been suggested that the regularity is due to a completely random distribution of the component molecules resulting from thermal agitation, such as exists in ideal solutions. Such randomness cannot, of course, persist to indefinitely low temperatures in the presence of orienting or associating forces or unequal molecular fields. On lowering the temperature to the critical solution temperature, for example, the two species begin to unmix, giving a very different statistical distribution, and submicroscopic aggregates certainly form above this point; however, the range of temperature and concentration in which this aggregation occurs to any appreciable extent is certainly very small, as shown by theoretical and experimental studies of opalescence in this region.[2] Where this randomness exists, it suggests, because of the well-known connection between entropy and probability, that the entropy of mixing for a regular solution should be the same as for an ideal solution of the same mole fraction. Writing the partial molal entropies of component X_1 in a regular and in an ideal solution as $\bar{S}_1^r$, and $\bar{S}_1^i$, respectively, the above conclusion can be expressed by the equation,

$$\bar{S}_1^r - \bar{S}_1^i = 0. \qquad (1)$$

Since the corresponding partial molal free energies are related to the entropies by the equation, $\partial(\bar{F}_1^r - \bar{F}_1^i)/\partial T = -(\bar{S}_1^r - \bar{S}_1^i)$, we can integrate, and write

$$\bar{F}_1^r - \bar{F}_1^i = f(N), \qquad (2)$$

where N denotes mole fraction. The form of $f(N)$ is the subject of this paper. It will be approached through the consideration of the energy of mixing. This is readily possible for regular solutions, since, when $\Delta S = 0$, $\Delta F = \Delta H$, and since, further, $\Delta H = \Delta E$ when $\Delta(PV)$ is negligible.

In order for a solution to be ideal, the potential of a given molecule must not alter when one species among the surrounding molecules is replaced by another. A solution will be regular even if this molecular potential varies, provided that the randomness of arrangement is still "ideal," i.e., complete. This can hardly be the case, in

[1] J. H. Hildebrand, J. Am. Chem. Soc. **51**, 66 (1929); also Proc. Nat. Acad. Sci. **13**, 267 (1927).

[2] Cf. M. v. Schmoluchowski, Ann. d. Physik [4] **25**, 205 (1908); A. Einstein, ibid. **33**, 1275 (1910); Wo. Ostwald, ibid. **36**, 848 (1911).

general, unless the molecules are symmetrical and nonpolar. (Isolated cases may be cited where a change of species, even with polar molecules, would probably not materially alter the arrangement, as with two isomeric esters.) We will assume equal molecular volumes, which allows a more satisfactory statistical treatment, before taking up the more difficult case of unequal volumes.

The form of $f(N)$ has been deduced by van Laar,[3] from van der Waals theory, and by Heitler[4] by aid of the assumption that the liquid has a cubic lattice structure. The former treatment shares the unconvincing character of the van der Waals equation, and the latter depends upon a picture which cannot be true. Heitler failed, moreover, as shown by his chosen examples, to recognize the essential limitation to systems that we have called regular. The following derivation is, nevertheless, to be regarded as a refinement of his method.

The nature of liquid structure has been revealed by the studies of x-ray scattering made recently by H. Menke[5] in Debye's laboratory.

The structure of a liquid is expressed by a probability function, W (designed W_r by Menke; the subscript is here omitted for the sake of simplicity) which is defined as follows. The probability that an element dV of a volume of liquid, V, will contain the center of a given molecule (e.g., the nucleus of a central atom) is dV/V. The probability that a given pair of molecules will occupy two such elements is $(dV/V)^2$, provided they are sufficiently far apart that they do not interfere with each other. When, however, the two elements of volume are very small, and their distance apart, r, is of the same order of magnitude as the molecular radius, the volume of the molecules themselves will help to determine their positions, and the probability must be expressed by $W(dV/V)^2$, where W is a function of r. Fig. 1A gives the plot of this function for mercury as determined by Menke from x-ray scattering.

It is possible to use W to calculate the number

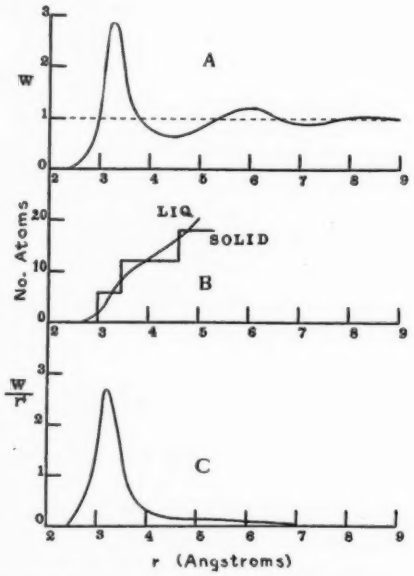

FIG. 1. A, Menke probability function for mercury; B, number of atoms within distance r of a central atom, for liquid and solid; C, potential function, area under which indicates small contribution to heat of vaporization of molecules beyond the "first layer."

of molecules (atoms in the case of mercury) surrounding a central one at various distances. The number of nuclei per cc in the liquid as a whole is N/v, where N is the Avogadro number and v the molal volume. The number of molecules in a large sphere of liquid would of course be $(N/v)(4/3)\pi r^3$. In a small sphere, however, the number surrounding a central one is

$$(N/v)\int_0^r 4\pi r^2 W dr.$$

Fig. 1B shows this number for various values of r calculated by graphic integration. The distribution for solid mercury is given for comparison (without correcting for the difference in density) and shows very strikingly how far the structure of the liquid differs from that of the solid.

The number in a spherical shell of radius r and thickness dr is $(4\pi N/v)Wr^2dr$. The number of molecules that can be regarded, in succession, as "central" to a shell is N, hence the total number of pairs at distance r from each other is $(2\pi N^2/v)Wr^2dr$ (the numerical coefficient is changed from 4 to 2 to avoid counting each pair twice). If the potential of each pair is φ (a function of r) the potential of all the pairs at distance r is

[3] J. J. van Laar, Zeits. f. physik. Chemie **A137**, 421 (1928).

[4] Heitler, Ann. d. Physik [4] **80**, 630 (1928). Cf. also J. H. Hildebrand and E. J. Salstrom, J. Am. Chem. Soc. **54**, 4257 (1932).

[5] (a) H. Menke, Phys. Zeits. **33**, 593 (1932); cf. further, (b) O. Kratky, ibid. **34**, 482 (1933).

$(2\pi N^2/\mathrm{v})\varphi W r^2 dr$, and the potential energy of the whole liquid, per mole, is

$$\mathrm{E}^{0\prime} = (2\pi N^2/\mathrm{v}) \int_0^\infty \varphi W r^2 dr. \tag{3}$$

If we assume with London,[6] and others, that the attraction potential between molecules is given by $-k/r^6$, and neglect the repulsive potential, the potential energy of a mole of liquid would be

$$\mathrm{E}^{0\prime} = (-2\pi N^2 k/\mathrm{v}) \int_0^\infty (W/r^4) dr. \tag{4}$$

Curve C, Fig. 1, shows the curve of W/r^4 plotted against r, the area under which represents the above integral. It is here given to illustrate how little is contributed by molecules beyond those immediately adjacent.

We will next consider a solution of n_1+n_2 moles of species X_1 and X_2, whose molal volumes are identical, v, and which mix without volume change. The number of molecular centers of both species per cc is the same as for the pure substance, i.e., N/v; the probability function should be the same for both species, hence the total number of molecules of both kinds at distance r from a central molecule is $(4\pi N/\mathrm{v}) W r^2 dr$, and since there are $(n_1+n_2)N$ possible central molecules, the total number of pairs at distance r from each other is $(2\pi N^2/\mathrm{v})(n_1+n_2) W r^2 dr$. The chances that these pairs are of the respective species X_1X_1, X_2X_2 and X_1X_2 are $n_1^2/(n_1+n_2)^2$, $n_2^2/(n_1+n_2)^2$ and $2n_1n_2/(n_1+n_2)^2$, hence the number of pairs of each kind are, respectively

$$\frac{2\pi N^2}{\mathrm{v}} \cdot \frac{n_1^2}{n_1+n_2} \cdot W r^2 dr, \quad \frac{2\pi N^2}{\mathrm{v}} \cdot \frac{n_2^2}{n_1+n_2} \cdot W r^2 dr$$

and

$$\frac{4\pi N^2}{\mathrm{v}} \cdot \frac{n_1 n_2}{n_1+n_2} \cdot W r^2 dr.$$

The potentials of single pairs are functions of r of the same form, designated by φ_{11}, φ_{22} and φ_{12}, where the subscripts denote the species forming the pair. The potential of all the X_1X_1 pairs at distance r is then $(2\pi N^2/\mathrm{v}) \cdot [n_1^2/(n_1+n_2)] \cdot \varphi_{11} W r^2 dr$, and the potential of all X_1X_1 pairs throughout the liquid is $(2\pi N^2/\mathrm{v}) \cdot [n_1^2/(n_1+n_2)] \cdot \int_0^\infty \varphi_{11} W r^2 dr$, and similarly for the X_2X_2 and X_1X_2 pairs. Adding these gives the potential energy of the entire liquid of n_1+n_2 moles, designated by E'. The partial molal energy of X_1 in this solution is $\bar{\mathrm{E}}_1 = (\partial E'/\partial n_1)_{n_2}$, which gives

$$\bar{\mathrm{E}}_1' = \frac{2\pi N^2}{\mathrm{v}} \cdot \frac{1}{(n_1+n_2)^2} \Big[(n_1^2 + 2n_1n_2) \times \int_0^\infty \varphi_{11} W r^2 dr - n_2^2 \int_0^\infty \varphi_{22} W r^2 dr + 2n_2^2 \int_0^\infty \varphi_{12} W r^2 dr \Big]. \tag{5}$$

Subtracting $\mathrm{E}_1^{0\prime}$ for a mole of the pure X_1, setting $\bar{\mathrm{E}}_1' - \mathrm{E}_1^{0\prime} = \bar{\mathrm{E}}_1 - \mathrm{E}_1^0$, the difference in potential energy equal to the difference in total energy, and collecting yields

$$\bar{\mathrm{E}}_1 - \bar{\mathrm{E}}_1^0 = (2\pi N^2 \mathrm{N}_2^2/\mathrm{v}) \Big[2 \int_0^\infty \varphi_{12} W r^2 dr - \int_0^\infty \varphi_{11} W r^2 dr - \int_0^\infty \varphi_{22} W r^2 dr \Big]. \tag{6}$$

If we set $\varphi_{11} = -k_{11}/r^6$, etc., we obtain an approximation formula

$$\bar{\mathrm{E}}_1 - \mathrm{E}_1^0 = \frac{2\pi N^2 \mathrm{N}_2^2}{\mathrm{v}} (k_{11} + k_{22} - 2k_{12}) \int_0^\infty \frac{W}{r^4} dr. \tag{7}$$

If, as has been frequently assumed, $k_{12} = (k_{11}k_{22})^{\frac{1}{2}}$, then

$$k_{11} + k_{22} - 2k_{12} = [(k_{11})^{\frac{1}{2}} - (k_{22})^{\frac{1}{2}}]^2. \tag{8}$$

By setting up the equations in the form of Eq. (3) for both pure components it can be seen that Eq. (7) can be transformed into

$$\bar{\mathrm{E}}_1 - \mathrm{E}_1^0 = [(\mathrm{E}_1^0)^{\frac{1}{2}} - (\mathrm{E}_2^0)^{\frac{1}{2}}]^2 \mathrm{N}_2^2. \tag{9}$$

This, unlike 6, can hardly be a rigid equation on account of the additional assumptions involved, especially Eq. (8), and it is not our purpose to subject it to a searching experimental test at this juncture. One illustration only will be given. Iodine and carbon disulfide have nearly identical molal volumes, 59.0 cc and 60.7 cc, respectively, at 25°C. Both have zero dipole moment. The solubility of iodine in carbon disulfide agrees closely[1] with the equation $\ln \mathrm{N}_1^i/\mathrm{N}_1 = b\mathrm{N}_2^2/RT$,

[6] F. London, Zeits. f. Physik **63**, 245 (1930).

where b is a constant. This corresponds to the relation $\bar{F}_1^r - \bar{F}_1^i = bN_2^2$. The value of b calculated from solubility is 870 cal. The energies of vaporization at 25° are, respectively, 11,300 cal. and 5970 cal., from which we calculate $[(E_1^0)^{\frac{1}{2}} - (E_2^0)^{\frac{1}{2}}]^2 = 840$ cal., a sufficiently good agreement.

Eq. (6), at any rate, should correspond well for systems conforming to the conditions postulated. The function, W, should depend chiefly on the molecular radius, for symmetrical molecules, and be relatively independent of the species of molecule. It should flatten out as the temperature increases and the departure from hexagonal close packing[5b] becomes greater.

Let us now consider a solution of two species whose molecular volumes are different, and where a single common probability function can not suffice. Consider a spherical shell of thickness dr, at a distance r, from the center of a central molecule. The volume of this shell is $dV = 4\pi r^2 dr$. The number of X_1 molecules whose centers are in 1 cc of liquid is $n_1 N/V$, where V is the total volume of the solution. The shell under consideration would contain $n_1 N dV/V$ molecular centers of X_1 if there were no interference due to the size of the central molecule, but since such interference exists a probability function will have to be introduced for each combination of species. This will be designated by subscripts, W_{11} for the X_1X_1 combination, etc.

The total number of molecules in the solution, each able to play the rôle of "central" molecule, is $(n_1+n_2)N$. The number of central molecules of X_1 is $n_1 N$. The number of X_1 molecules in the shell surrounding each is $n_1 N W_{11} dV/V$. The number of X_1X_1 "pairs" throughout the solution at distance r is $\frac{1}{2} n_1^2 N^2 W_{11} dV/V$. (The $\frac{1}{2}$ is introduced to avoid counting each pair twice.) The potential of each X_1X_1 pair is φ_{11}, hence the potential of all such pairs at distance r is $(2\pi n_1^2 N^2/V) \cdot \varphi_{11} W_{11} r^2$, and the potential of all X_1X_1 pairs throughout the liquid at all distances, is $(2\pi N^2 n_1^2/V) \int_0^\infty \varphi_{11} W_{11} r^2 dr$. We will abbreviate this expression by writing for the integral the symbol $\int(11)$, giving $(2\pi N^2 n_1^2/V) \int(11)$. For the X_2X_2 we get the corresponding expression, $(2N^2 n_2^2/V) \int$ (22) and for the X_1X_2 and X_2X_1 pairs, which are identical, $(4\pi N^2 n_1 n_2/V) \int$ (12).

We will pause here to consider the probability functions, W_{11}, W_{22}, W_{12} and W_{21}. It would seem that the position of the first maximum, as shown in Fig. 1A, should be the same for the solution as for the pure substance, since it depends chiefly on the sum of the two molecular radii, increased somewhat by thermal agitation. The position of the second maximum would, however, be different, for, if X_2 is thought of as smaller than X_1, two molecules of X_1 might be separated by an X_2 instead of another X_1. The frequency of this separation would depend upon the proportion of X_2 molecules, hence W_{11} for the solution would be a function of the composition beyond the first maximum. A quantitative formulation of this dependence seems very difficult and we shall, for the present, take refuge in the fact that the potential is but little affected by the course of W much beyond the first maximum, as illustrated by the area under the curve in Fig. 1C, permitting us to set W_{11} for solution and pure X_1 equal to each other as a first approximation. We are undertaking an experimental study of these functions.

We should note, further, that although W_{12} and W_{21} might appear to be different, the number of X_1X_2 and X_2X_1 pairs is the same, hence $\int(12) = \int(21)$. We are now prepared to write for the total potential energy of n_1+n_2 moles of solution,

$$E' = (2\pi N^2/V)\left[n_1^2 \int(11) + n_2^2 \int(22) + 2n_1 n_2 \int(12)\right]. \quad (10)$$

The partial molal energy of X_1, i.e., $\bar{E}_1'$, is obtained, as before by differentiating with respect to n_1 at constant n_2, remembering that $V = n_1 v_1 + n_2 v_2$, approximately. We should write, more strictly, $V = n_1 \bar{v}_1 + n_2 \bar{v}_2$, where $\bar{v}_1$ and $\bar{v}_2$ are the partial molal volumes, but this could hardly be worth the added labor in view of the other approximations involved. Carrying out this differentiation and subtracting the value of $E_1^{0\prime}$ from Eq. (3) gives

$$\bar{E}_1 - E_1^0 = 2\pi N^2 \left(\frac{n_2 v_2}{n_1 v_1 + n_2 v_2}\right)^2 \cdot v_1 \left[\frac{2}{v_1 v_2} \int(12) - \frac{1}{v_1^2} \int(11) - \frac{1}{v_2^2} \int(22)\right]. \quad (11)$$

We may regard this equation as valid for any system of the sort here considered. To apply it rigidly will, however, require more data than we now have concerning probability functions and intermolecular potentials. Indeed, one of its principal uses is likely to be the throwing of light upon the latter question. In the meantime we may derive from it an approximation formula which may be tested by the aid of solubility data now at hand.

First, we will neglect the repulsive potential and write for the attractive potential simply $=-k/r^6$. This may make an appreciable error in dealing with mercury at ordinary temperatures,[7] which is not much expanded over its volume at 0°K. Fig. 1B shows the average number of molecules in the liquid which are closer together than in the solid. For substances at temperatures nearer their boiling points the effect of repulsion would be less.

Next, we will assume, by analogy with Eq. (8), that

$$\int(12)=\int^{\frac{1}{2}}(11)\times\int^{\frac{1}{2}}(22). \qquad (12)$$

The physical significance of this is as follows. Since the form of the function, W, is the result of a thermally disturbed close packed arrangement, it may be assumed to be, for all spherical molecules, approximately the same function of the molecular domain. This could hardly be strictly true, for even the curve for a pure liquid will surely flatten with increasing thermal agitation, hence curves for different liquids must be, in general, slightly different. However, if we neglect differences in flattening as an effect of higher order, we may express the different W's as the same function of the position of the first maximum, $\mathbf{r}$, for each, i.e., $W=f(r/\mathbf{r})$. The curves obtained by Menke for mercury and gallium, when plotted against $r/\mathbf{r}$, agree very closely, probably within the limit of error. We will therefore assume that the integral $k\int_0^\infty (W/r^4)dr$, if transformed into

$$\frac{k}{\mathbf{r}^3}\int_0^\infty\left(\frac{\mathbf{r}}{r}\right)^4 f\left(\frac{r}{\mathbf{r}}\right)d\left(\frac{r}{\mathbf{r}}\right),$$

would be the same for all substances so far as the integrand is concerned. Introducing this, with subscripts appropriate to each combination, into Eq. (12), and cancelling out the integrand, leaves

$$k_{12}/\mathbf{r}_{12}^3=(k_{11}k_{22}/\mathbf{r}_{11}^3\mathbf{r}_{22}^3)^{\frac{1}{2}}. \qquad (13)$$

If we write $\mathbf{r}_{12}=\frac{1}{2}(\mathbf{r}_{11}+\mathbf{r}_{22})$, which should be very nearly true, Eq. (13) becomes

$$k_{12}=\frac{(\mathbf{r}_{11}+\mathbf{r}_{22})^3}{8\mathbf{r}_{11}^3\mathbf{r}_{22}^3}(k_{11}k_{22})^{\frac{1}{2}}. \qquad (14)$$

The coefficient of $k_{11}k_{22}$ does not differ much from unity, even when $\mathbf{r}_{11}$ and $\mathbf{r}_{22}$ differ considerably; for example, it is only 1.06 when $\mathbf{r}_{11}$ and $\mathbf{r}_{22}$ differ by 50 percent. Eq. (14) is therefore nearly identical with Eq. (8), which is generally regarded as approximately true. We are therefore encouraged to try the effect of combining Eq. (12) with Eq. (11), and then combine the result with Eq. (3) applied to each pure component. The final result is

$$\bar{E}_1-E_1^0=\left(\frac{n_2V_2}{n_1V_1+n_2V_2}\right)^2 V_1\left[\left(\frac{E_1^0}{V_1}\right)^{\frac{1}{2}}-\left(\frac{E_2^0}{V_2}\right)^{\frac{1}{2}}\right]^2. \qquad (15)$$

The equation derived by van Laar[3] with the aid of the van der Waals equation reduces to this under the conditions herein postulated. The same equation has been given by Scatchard[8] by the aid of some of the same assumptions as here employed, plus the arbitrary one that the cohesive energy of a mole of mixture is given by the expression

$$\frac{a_{11}V_1^2n_1^2+2a_{12}V_1V_2n_1n_2+a_{22}V_2^2n_2^2}{(n_1V_1+n_2V_2)^2},$$

where the a's are regarded as constants (equivalent to the van der Waals constants.)

Experimental data now being obtained in this laboratory upon tetrahalides and their solutions will enable us to make a more searching test of the laws of regular solutions than is now possible. In the meantime we may cite certain iodine solutions as most nearly conforming to the requirements. We have data for the solubility of iodine[9]

[7] J. H. Hildebrand, Phys. Rev. **34**, 984 (1929).

[8] Geo. Scatchard, Chem. Rev. **8**, 321 (1931).

[9] J. H. Hildebrand and C. A. Jenks, J. Am. Chem. Soc. **42**, 2080 (1920); G. R. Negishi, L. H. Donnally and J. H. Hildebrand, ibid. Nov. (1933); cf. also reference 1.

in the tetrachlorides of carbon, silicon and titanium, all very symmetrical molecules with zero electric moment. To these we shall add the solutions in carbon disulfide. Table I gives values of $\bar{F}_2^r - \bar{F}_2^i$ for iodine, calculated from the experimental data by the expression $RT \ln (N_2^i/N_2)$, in comparison with values of $\bar{E}_2^r - E_2^0$, calculated from Eq. (12). The values of v_1, v_2, E_1^0 and E_2^0 used are indicated in Table I. They are taken from the usual sources which need not here be elaborated, since the purpose is merely illustrative. The values are given for 25° only, since the uncertainty regarding the change of the heat of vaporization with temperature hardly makes it worth while to attempt to check the small temperature trend.

The agreement between the figures in the last two columns may be regarded as very satisfactory, indicating that the solubility of iodine could be calculated by the aid of Eq. (12) without serious error.

TABLE I. *Test of Eq. (12) by means of solubility data for iodine at 25°C.*

Solvent	v cc per mole	E° cal.	Mole percent I_2 in soln.	$\bar{F}_2 - \bar{F}_2^i$ cal.	$\bar{E}_2 - E_2^0$ cal.
$TiCl_4$	110.5	8370	2.150	1365	1535
$SiCl_4$	115.4	6610	0.499	2220	2320
CCl_4	97.1	7088	1.594	1725	1635
CS_2	60.7	6070	5.76	770	815
Iodine	59.0	11330			

The solubility data for sulfur and for stannic iodide might also be used, but uncertainties regarding heats of vaporization at the temperatures in question diminish the significance of the comparisons, and they are here omitted.

One of the most important uses for the equations here derived will doubtless be to throw light upon the important question of the potential functions for pairs of like and unlike molecules. We expect to be able to report upon this in the near future.

Technique for the Electrolytic Production of H^2H^2O

HUGH S. TAYLOR, HENRY EYRING AND ARTHUR A. FROST, *Frick Chemical Laboratory, Princeton, N. J.*

(Received October 2, 1933)

A detailed description of a method of preparing heavy water in quantity is given and yields obtained over a period of months reported.

NUMEROUS requests for a detailed description of our methods for the production of water rich in heavy hydrogen isotope, first outlined elsewhere,[1] prompt the following more detailed discussion of our present practice in the light of experience attained during the past summer in the treatment of 600 gallons of used electrolyte with the production of some 80 cc of water containing upwards of 90 percent heavy hydrogen.

ELECTROLYSIS WITHOUT RECOVERY

Electrolyte from commercial electrolytic cells is distilled in order to remove the large amounts of carbonate and hydroxide present. Sodium hydroxide is added to make the distillate approximately 0.5 molar. This solution is electrolyzed in a battery of 210 cells, made from hydrometer jars, 4 cm diameter and 25 cm high, each of approximately 200 cc capacity, water-cooled by immersion in a tank containing flowing water and carrying a current of six to seven amperes. The electrodes are made of nickel plate, a strip 60 cm by 3 cm being bent twice at right angles so as to form anode and cathode in neighboring cells. The battery is subdivided into units of 30 cells each, the potential of the direct current supplied from a motor generator being adjusted so as to yield the amperage employed. After three days, when the electrolyte has diminished to 1/6 or 1/7th its original volume, it is replaced by a fresh batch of electrolyte. This process is so conducted that, while one cell is being emptied with the aid of suction, a neighboring cell is being filled with fresh electrolyte by a siphon, the whole process requiring about one hour. The concentrated electrolyte from the first electrolysis is partially neutralized by bubbling carbon dioxide through it so as to form the carbonate, with fuchsin as an indicator, and then distilled. This water is now added to another group of cells containing water of the same grade but still containing all of the sodium hydroxide originally added with the water at the start of the first electrolysis. This procedure obviates setting aside a portion of the concentrated liquor before distillation which would later be added to the distillate to give the 0.5 molar solution for the second electrolysis. In this way three successive electrolyses are carried out resulting in water of density $D_{4°}^{20°} = 1.001$, approximately, and which therefore contains about 2.5 percent of the heavy hydrogen isotope.

ELECTROLYSIS WITH RECOVERY

From this stage onwards the hydrogen evolved during electrolysis contains an appreciable amount of the heavy isotope which should be recovered. This has been effected in the following way. Water containing sodium hydroxide is electrolyzed in cells of 200 cc capacity carrying a current of 10 amperes each and cooled in a trough of running water. The mixed electrolytic gas passes successively, as shown in Fig. 1,

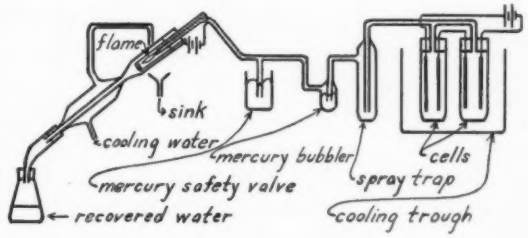

FIG. 1. Electrolytic concentration apparatus.

[1] H. S. Taylor and H. Eyring, Proc. Am. Phil. Soc. **72**, 255 (1933).

through a spray-trap, a mercury bubbler, through a T-tube immersed in mercury to form a safety valve for releasing excessive pressure, and finally flows out through a capillary nozzle of Pyrex glass where it burns as a flame. The water formed by the combustion is condensed in an inclined quartz tube surrounded at the lower end by a water jacket. A red-hot platinum filament is placed just beyond the flame in order to re-ignite the gas in case the flame should go out, which may happen every few minutes if the size of the capillary nozzle is not perfectly adjusted for the rate of gas flow from the cells. For a given rate of gas flow, if the capillary is too fine it is easily clogged up, presumably by alkali from the cells; on the other hand, if the capillary is too wide, the flame strikes back through the tubing but is prevented from reaching the cells through the mercury obstruction (about 1 cm) in the mercury bubbler. If the condenser tube in the neighborhood of the flame becomes too hot, the flame jumps to the platinum wire where it burns incompletely. This is prevented by cooling the tube slightly at this point, a fact that indicates the effect is due to an excessive pressure of water vapor which is known to hinder the explosive reaction. For two cells each carrying 10 amperes the optimum size for the capillary has been found to be about 0.3 mm. Each nozzle has a lifetime of about one day but is easily replaced by a new nozzle. Practically a quantitative recovery of the water can be obtained by this method.

Yield

In electrolyzing a dilute solution of the heavy water down to about 1/6th of the original volume, it is found that the concentration of the heavy hydrogen in the heavier 1/6th has increased by a factor of about four while the concentration in the lighter 5/6ths is about 2/5ths that of the original. The recovered water is therefore added to the water one stage back in the process. With the present arrangements seven successive stages suffice to produce water which contains about 99 percent of the hydrogen in the heavy form. Table I indicates the quality of the water now obtained at each stage of the process.

Table I.

Water obtained from electrolysis No.	Density D_4^{20}	% hydrogen which is heavy
I	0.998	—
II	0.999	0.5
III	1.001	2.5
IV	1.007	8
V	1.031	30
VI	1.098	93
VII	1.104	99

Table II shows the amounts of water put through each stage of the process during the period May 9th to September 27th, 1933.

Table II.

Electrolysis No.	Amt. electrolyzed May 9–Sept. 27, 1933
I	610 gal. (of commercial electrolyte)
II	90 gal.
III	52 l
IV	10.15 l
V	2.00 l
VI	420 cc
VII	82 cc

It is found that about 1 cc of water, 95 percent heavy, is obtained for every seven gallons of starting material. Assuming that the commercial electrolyte liquor contains one part in 3000 heavy hydrogen, it is easily calculated that approximately 10 percent of the heavy isotope of hydrogen present in the original starting material is obtained in the heavy water. Most of the loss occurs in the first three electrolyses where no attempt is made to recover the heavy hydrogen in the gas evolved during electrolysis.

The Ground State of the Hydrogen Molecule[1]

HUBERT M. JAMES AND ALBERT SPRAGUE COOLIDGE, *Harvard University*

(Received October 3, 1933)

The method used by Hylleraas in treating the He atom has been extended to the H_2 molecule. The method consists of setting up a wave function as a series in the five variables required, electronic separation being introduced explicitly as one of the variables. The coefficients are then determined so as to produce the lowest energy. The energy found is within 0.03 v.e. of the most probable experimental value, while the form and location of the potential energy curve for various internuclear distances agree with those deduced from spectra to within similar limits. The value of the function is computed for several configurations of the electrons, and compared with other approximations. Application of the method to other problems is discussed. A method is given for the numerical solution of secular equations of high degree.

INTRODUCTION

IN constructing an approximate wave function for the hydrogen molecule in its normal state (regarding the nuclei as fixed), one may follow several lines of attack. Heitler and London make use of the unchanged functions of the individual atoms, regarding the interatomic forces as small perturbations. In a molecule as close-coupled as H_2 this implied persistence of the identity of the individual atoms is quite unjustified, and the resulting wave function is a very poor approximation. In particular, it is inadequate in that it is a function of only four electronic coordinates, instead of the required five. Aside from convenience, the only advantage of the H–L scheme seems to be that it yields a result in which the energy of the molecule is represented as a sum of terms, some of which are just the energy of the separated normal atoms, so that these may be cancelled out and the remaining terms called "binding energy."

Several attempts have been made to improve the H–L method by grafting upon it the variation principle; the original atomic wave functions are modified by the introduction of one or two arbitrary parameters, which are then determined so as to give the lowest energy. Such schemes do give somewhat better results, but do not escape the essential shortcomings of the original method —the implication of atomic individuality, and the omission of an essential coordinate. Indeed, they actually sacrifice the peculiar advantage of the pure perturbation method, since the unchanged atomic energy no longer appears in the result.

In view of the theoretical weakness and practical inadequacy of these methods, it seemed to us worth while to explore rather carefully a third possibility. Abandoning all concern with individual atoms, we have tried to build up from the ground a suitable molecular wave function, containing the full number of required coordinates. While guided by certain considerations of limiting forms of this function, we have placed our main reliance on the variation principle, introducing a large number of parameters to be determined by its use.

DISCUSSION OF METHOD

Concerning the true function, we know the following: The ground state being a Σ state, the function must have rotational symmetry about the nuclear axis; that is, it may contain only the difference between the azimuthal angles of the two electrons, and not either angle alone. It must be symmetrical in the two electrons and also in the nuclei. In those parts of phase-space corresponding to one electron at a large distance from the nuclei, the wave function should approximate the product of a H-like function for the distant electron and an H_2^+-like function for the close electron. The last two conditions

[1] A preliminary notice appeared in Phys. Rev. **43**, 588 (1933). Some minor numerical inaccuracies in that notice have been corrected in the present paper.

suggest the use of elliptical coordinates and the insertion of an exponential factor reducing to e^{-r} for either electron at large distances. The necessary flexibility can best be secured by expressing the rest of the function as a power series in the five variables, with coefficients to be determined by variation to produce the minimum energy. Such coefficients can be handled much more satisfactorily than parameters appearing in other ways (for example, in exponents). Finally, Hylleraas,[2] working on another two-electron problem, the helium atom, has shown the advantage of taking the interelectronic distance itself, rather than the difference in azimuthal angle, as one of the independent variables.

Let the distances between the several particles involved, expressed in terms of the first Bohr radius, carry the usual symbols R, r_{1a}, r_{1b}, r_{2a}, r_{2b}, r_{12}. Then our coordinates are

$$\lambda_1=(r_{1a}+r_{1b})/R, \qquad \lambda_2=(r_{2a}+r_{2b})/R,$$

$$\mu_1=(r_{1a}-r_{1b})/R, \qquad \mu_2=(r_{2a}-r_{2b})/R,$$

$$\rho=2r_{12}/R,$$

and our trial function is

$$\psi=\Sigma_{mnjkp}C_{mnjkp}[mnjkp],$$

where $[mnjkp]$ stands for

$$(1/2\pi)e^{-\delta(\lambda_1+\lambda_2)}(\lambda_1{}^m\lambda_2{}^n\mu_1{}^j\mu_2{}^k\rho^p+\lambda_1{}^n\lambda_2{}^m\mu_1{}^k\mu_2{}^j\rho^p).$$

The summation is to extend over positive or zero values of the indices, subject to the restriction required by nuclear symmetry that $j+k$ must be even, and taking as many terms as shall prove necessary to give an acceptable approximation for the energy. This approximation will be found by computing the minimum value of $\int\int\psi^*H\psi dV_1dV_2$ which can be obtained by suitable choice of the coefficients C, subject always to the normalizing condition $\int\int\psi^*\psi dV_1dV_2=1$, and must always lie above the true energy.[3] The same choice of C's will then ensure that the trial function is as nearly as possible a solution of the wave equation $H\psi=E\psi$. The errors in the function itself will greatly exceed that in the energy; Eckart[4] has shown that if the errors are small, the former will be of the order of the square root of the latter.

In accordance with what has been said about the limiting form of the function, the exponent δ should have the value $R/2$, but additional flexibility can be secured without much additional calculation by regarding it as another arbitrary parameter to be varied. It should be noted that the effect of such variation upon the wave function will be slight (at least in the region where the function is important), provided that the coefficients are allowed to re-adjust themselves. For a small change in δ is equivalent to multiplication by a rapidly converging power-series in λ_1 and λ_2, the effect of which can be almost completely absorbed by appropriate changes in the coefficients of the smaller powers. Now, the whole computation must be done from the ground up for each new value of δ, while different values of R can be introduced with little trouble. Let us limit ourselves to a definite number, s, of terms in the series for ψ. Then for each value of δ there will be one R for which the best energy will be lower than it would have been for the same R if we had used any other δ. To find this, we must plot a series of curves (similar in appearance to Morse curves), each giving the energy computed with a given δ for various R's. The envelope of this family of curves will then show the lowest energy obtainable for any R with the corresponding most appropriate δ, and each δ used in computing will be the best possible for that particular R at which its curve osculates the envelope. (It turns out that δ should somewhat exceed $R/2$.)

For a given δ and R, it can be readily[3] shown that the values of the C's which minimize the computed energy are those which satisfy the system of equations

$$\begin{aligned}
&(H_{11}-\lambda S_{11})C_1+(H_{12}-\lambda S_{12})C_2+\cdots\\
&\qquad\qquad+(H_{1s}-\lambda S_{1s})C_s=0\\
&(H_{12}-\lambda S_{12})C_1+(H_{22}-\lambda S_{22})C_2+\cdots\\
&\qquad\qquad+(H_{2s}-\lambda S_{2s})C_s=0\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&(H_{1s}-\lambda S_{1s})C_1+(H_{2s}-\lambda S_{2s})C_2+\cdots\\
&\qquad\qquad+(H_{ss}-\lambda S_{ss})C_s=0.
\end{aligned}$$

[2] E. A. Hylleraas, Zeits. f. Physik **54**, 347 (1929).

[3] See, for instance, J. K. L. MacDonald, Phys. Rev. **43**, 830 (1933).

[4] C. Eckart, Phys. Rev. **36**, 878 (1930).

Here λ is a Lagrangian multiplier, and the condition of compatibility is the secular equation

$$\begin{vmatrix} H_{11}-\lambda S_{11} & H_{12}-\lambda S_{12} & \cdots & H_{1s}-\lambda S_{1s} \\ H_{12}-\lambda S_{12} & H_{22}-\lambda S_{22} & \cdots & H_{2s}-\lambda S_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ H_{1s}-\lambda S_{1s} & H_{2s}-\lambda S_{2s} & \cdots & H_{ss}-\lambda S_{ss} \end{vmatrix} = 0.$$

The H_{fg} and the S_{fg} are the matrix components of the Hamiltonian energy operator and of unity, respectively, between the fth and gth terms in the series, considered as numbered consecutively from 1 to s. Details of their computation, of the solution of the secular equation, and of the determination of the C's will be reserved for the Appendix. As usual, the lowest of the s values of λ turns out to be the required minimum energy, and can be found without determining the C's.

By way of finding out how many and which terms in the series would have to be used in order to get a good approximation for the wave function, we confined ourselves in the beginning to the equilibrium distance, $R=1.4$, and to $\delta=0.75$. It soon became apparent that the first few terms alone were capable of giving a far better energy value than any previously reported, and that incorporation of additional terms produced only rapidly diminishing improvements. In some cases, even terms involving low powers of the variables made such slight changes as to warrant leaving them out. Of course, the incorporation of each new term required the computation of its matrix elements with all the terms already in use and the solution of a new secular equation of higher degree. Strictly, a term should not be neglected until it had been tested in combination with all other terms. But it soon became clear that the improvement obtainable by any given term became progressively less important as the number of other terms present increased, so that we were justified in rejecting any term which at any stage in the building up of our function was found to produce a negligible improvement in the energy.

Results

An idea of the results obtained is given in Table I, in which the energy at several stages is compared with that given by other methods, and with the observed value. The last is obtained by adding 0.27 v.e., the zero-point energy of the normal molecule, to the heat of dissociation, which is 4.46 ± 0.04 according to Richardson and Davidson,[5] while Mulliken[6] gives 4.44. Much of the table is copied from Rosen,[7] who, however, seems to have neglected the zero-point energy in giving the observed value. The values in the last column are those for which the various functions (other than the series functions here presented) give their best results. They are sufficiently near $R=1.4$ for purposes of comparison, except the Heitler-London function, whose minimum energy lies at considerably greater R, so that it seems worth while to include also the energy given by this function at the true distance.

Table I.

Function	Total energy atomic units	Binding energy electron-volts	Internuclear distance, Bohr radii
One term	−2.189	−2.56	1.40
5 terms	−2.33290	−4.507	1.40
11 terms	−2.34580	−4.682	1.40
13 terms	−2.34693	−4.697	1.40
Without r_{12}	−2.3154	−4.27	1.40
Heitler-London	−2.21	−2.9	1.40
Heitler-London	−2.24	−3.2	1.51
Wang	−2.278	−3.76	1.42
Rosen	−2.297	−4.02	1.416
Observed		-4.73 ± 0.04	1.40

Table II.

Terms	Coefficients in normalized functions			
[00000]	1.69609	2.23779	2.29326	2.22350
[00020]		0.80483	1.19526	1.19279
[00110]		−0.27997	−0.49921	−0.45805
[10000]		−0.60985	−0.86693	−0.82767
[10200]			−0.13656	−0.17134
[10020]			−0.07214	−0.12101
[10110]			0.14330	0.12394
[20000]			0.06621	0.08323
[00001]		0.19917	0.33977	0.35076
[00021]				0.07090
[00111]			−0.03143	−0.01143
[10001]				−0.03987
[00002]			−0.02456	−0.01197

[5] O. W. Richardson and P. M. Davidson, Proc. Roy. Soc. **A123**, 466 (1929).

[6] R. S. Mulliken, Rev. Mod. Phys. **4**, 78 (1932).

[7] N. Rosen, Phys. Rev. **38**, 2099 (1931).

The terms used in the first four functions of Table I, together with their normalized coefficients, are given in Table II. It is interesting to note that the simple exponential wave function,

$$e^{-\delta(\lambda_1+\lambda_2)} \approx e^{-r_{1a}/2} \times e^{-r_{1b}/2} \times e^{-r_{2a}/2} \times e^{-r_{2b}/2}$$

already gives a binding energy comparable to that of the H–L function; indeed, by a suitable choice of δ the result can be made better, when for both functions $R=1.4$. In the energy thus calculated, there is nothing resembling the "exchange integrals" of the H–L treatment; this raises the question whether the importance of the "exchange terms," frequently assumed to represent the essential nature and magnitude of chemical binding, may not have been overemphasized. The five-term function is offered as a practical compromise between simplicity and accuracy. The reason for including both 11- and 13-term functions is really accidental. Due to a mistake, we at first believed that the contributions from the terms [00021] and [10001] could be neglected, and we therefore took the 11 terms as a basis for investigating the effects of varying δ and R (discussed later). The results showed that in fact we had hit upon the best possible values of both. In going over the work, we discovered the mistake, and found that a slight improvement could be obtained by including the two given terms, and we therefore offer the 13-terms function as the best known approximation to the true wave function of H_2. It did not seem worth while to repeat the variations of δ and R with this slightly different form of the function.

The energy given by the 13-term function is within the range fixed by experiment, but higher than the most probable value. It is beyond doubt that if still more terms were included, the result could be still further depressed. We have made a rather careful estimate of what could be gained in this way, reasoning by analogy from the contributions actually found from series of related terms; it seems safe to say that we have reached the limit of convergence within from 0.01 to 0.05 volt. The theoretical energy of the hydrogen bond is, then, 4.73 ± 0.02 volt, as compared with the experimental 4.73 ± 0.04 volt.

We found it possible to reach the value -4.27 with a combination of terms with $p=0$, and therefore not including r_{12} in the wave function. (The coefficients were not determined in this case.) This is substantially the same as the value reported by Hylleraas[8] as the best which he could obtain by an involved computation which he does not reproduce, and is probably close to the limit obtainable with wave functions in four coordinates.

It may be of interest to compare the magnitudes of some of the functions themselves, computed for various positions of the electrons, and to see how nearly they satisfy Schrödinger's equation. Table III shows values of ψ and of

TABLE III.

Coordinates						
r_{1a}	0.35	0.35	0.35	0.35	1.4	1.4
r_{1b}	1.4	1.4	1.4	1.4	1.4	1.4
r_{2a}	0.35	0.7	0.7	0.7	1.4	1.4
r_{2b}	1.75	0.7	1.4	1.4	1.4	1.4
r_{12}	0.525	0.7425	0.3556	1.0334	0	2.4249
$\varphi_1-\varphi_2$	—	—	0°	180°	0°	180°
Values of ψ (normalized)						
H–L	0.0532	0.0849	0.0533	0.0533	0.0219	0.0219
Rosen	0.0702	0.1215	0.0710	0.0710	0.0249	0.0249
5-term	0.0798	0.1172	0.0698	0.0777	0.0162	0.0271
11-term	0.0783	0.1180	0.0667	0.0777	0.0130	0.0271
13-term	0.0769	0.1162	0.0656	0.0772	0.0142	0.0269
Values of $(H-E)\psi$, in 27.08 volts						
H–L	−0.0425	−0.0481	0.0460	−0.0522	$0.0313/\rho-0.0052$	
Rosen	−0.0310	−0.0015	0.0879	−0.0429	$0.0356/\rho-0.0038$	
5-term	0.0148	0.0093	0.0653	−0.0097	$0.0102/\rho-0.0002$	
11-term	0.0117	0.0109	0.0356	−0.0060	$-0.0034/\rho-0.0005$	
13-term	0.0055	0.0060	0.0269	−0.0021	$0.0028/\rho-0.0009$	

$(H-E)\psi$, calculated in each case for $R=1.4$; the unit of energy is here (as in all our computations) 27.08 v.e., twice that adopted for purposes of comparison in Table I.

It is natural to regard the discrepancies between the various functions as a rough indication of their probable errors. The differences shown in Table III are representative of a large number of points which have been calculated. Evidently the H–L function is much too diffuse, giving ψ^2 only about half its proper value in the region of importance, and a corresponding excessive value in more remote regions in order to preserve normalization. Between the other functions no systematic differences are conspicuous except when we compare points in phase-space for which the elliptical coordinates of the two electrons are respectively the same, but the differences in azimuth, and hence the interelectronic separations, are different. Columns 3

[8] E. A. Hylleraas, Zeits. f. Physik **71**, 739 (1931).

and 4, 5 and 6, offer such comparisons. Here the present functions show their characteristic advantage of being able to assume smaller values for the smaller electronic separations.

The discrepancies are in general in good agreement with Eckart's[4] criterion, which can be thus stated: the root-mean-square error in ψ is equal to $(\epsilon/\Delta E)^{\frac{1}{2}}$, where ϵ is the error in the calculated energy, and ΔE represents a kind of average of the separation between the lowest energy level and the other levels of the same symmetry. According to Mulliken,[6] the second lowest $^1\Sigma_g^+$ state has an energy about 20 volts above the ground state (for $R=1.4$). The state of complete ionization of the fixed-center molecule would evidently have the positive absolute energy 19.3 volts (due to nuclear repulsion), or 51 volts above the ground state. If we take ΔE as 35 volts, ϵ as 1.8, 0.71, 0.22, 0.05, and 0.03 volts for the five functions, respectively, the mean errors should be 23 percent, 14 percent, 8 percent, 4 percent, and 3 percent. The errors in ψ^2, and presumably in quantities calculated by its aid, will be double these figures.

Corroboratory evidence is offered by the values of $(H-E)\psi$. This quantity, which vanishes for the correct function, seems on the basis of five computed points to tend to amount to something like 80 percent, 40 percent, 12 percent, 9 percent, 6 percent, of ψ. Reasoning similar to Eckart's (see Appendix) leads to the conclusion that the root-mean-square average of these ratios should be somewhat greater than $(\epsilon\times\Delta E)^{\frac{1}{2}}$, or 29 percent, 19 percent, 10 percent, 5 percent, 4 percent, respectively. The magnitude of $(H-E)\psi/\psi$ varies greatly from point to point, but at each point there is a clear tendency for the values to run proportional to the square-roots of the assumed energy errors, thus indicating that those errors are indeed of the right order of magnitude, and that the correct function would give the experimental value.

It is instructive to see what happens when the electrons coincide and $r_{12}=0$. Since the term $1/r_{12}$ in the potential energy becomes infinite, Schrödinger's equation requires either that ψ vanish or that a cancelling infinite term arise from the Laplacian or kinetic energy part of the operator H. The second alternative is evidently the correct one. In the fifth column of Table III, the entries under $(H-E)\psi$ show the coefficient of the $1/\rho$ term in $H\psi$, which outweighs all other terms as ρ becomes small. In the H–L and Rosen functions, this term comes solely from the potential energy; in the series functions, especially the 13-term function, cancellation by a term from the Laplacian is approximately complete. The function itself shows no sign of approaching zero as it is made more accurate.

In order to see whether still lower energy values could be obtained by varying δ and/or R, we computed a number of energies for different values of these quantities, using the same eleven terms in the series. As a basis of comparison, we took as "observed" values those given by a Morse curve constructed as recommended by Mulliken,

$$U(r)=D_e[1-e^{-a(r-r_e)}]^2; \qquad a=(2\pi^2\mu c\omega_e^2/hD_e)^{\frac{1}{2}}.$$

taking $D_e=4.73$ e.v., $\omega_e=4375\text{cm}^{-1}$, $r_e=1.40\,a_H$. $\mu=0.8309\times10^{-24}$ g.

TABLE IV.

R		1.2	1.3	1.4	1.5	1.6	1.7
Binding energy, electron-volts	$\delta=0.875$			−4.67	−4.63	−4.52	−4.35
	$\delta=0.75$	−4.41	−4.62	−4.68	−4.63	−4.49	
	Morse	−4.49	−4.67	−4.73	−4.68	−4.57	−4.40

The results are summarized in Table IV. The significance of these results will best be seen from the graph, Fig. 1. It will be noticed, first, that the choice already discussed, $\delta=0.75$, R

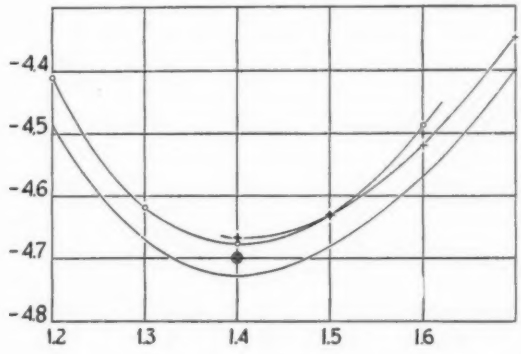

FIG. 1. Energy of H_2 in electron-volts, against separation of the nuclei, in Bohr radii. (Energy of dissociated molecule taken as zero.) Lowest curve is "experimental" Morse curve. Circle indicates point computed with $\delta=0.75$, 11-term function. Cross indicates point computed with $\delta=0.875$, 11-term function. Double circle indicates point computed with $\delta=0.75$, 13-term function.

TABLE V.

Terms	Coefficients							
	$\delta=0.75$				$\delta=0.875$			
	$R=1.2$	$R=1.3$	$R=1.4$	$R=1.5$	$R=1.4$	$R=1.5$	$R=1.6$	$R=1.7$
[00000]	2.64085	2.47282	2.29326	2.10831	2.55009	2.42055	2.28454	2.14343
[00020]	1.05889	1.14862	1.19526	1.20876	1.32769	1.46843	1.56080	1.61252
[00110]	−0.35865	−0.44227	−0.49921	−0.53385	−0.48136	−0.59679	−0.68500	−0.74964
[10000]	−0.75010	−0.82703	−0.86693	−0.87515	−0.76638	−0.84500	−0.89956	−0.92919
[10200]	−0.07001	−0.10926	−0.13656	−0.15443	−0.06026	−0.12001	−0.16569	−0.19941
[10020]	−0.00474	−0.04501	−0.07214	−0.08931	+0.03746	−0.02644	−0.07409	−0.10817
[10110]	0.05723	0.10936	0.14330	0.16388	0.03528	0.11560	0.17625	0.22060
[20000]	0.06082	0.06100	0.06621	0.07168	0.07933	0.07208	0.07421	0.07970
[00001]	0.46092	0.39769	0.33977	0.28909	0.54697	0.48577	0.42713	0.37287
[00111]	−0.01097	−0.02408	−0.03143	−0.03508	−0.00316	−0.02137	−0.03420	−0.04269
[00002]	−0.02616	−0.02596	−0.02456	−0.02264	−0.03099	−0.03198	−0.03135	−0.02974

$=1.4$, gives the lowest energy of any calculated. Second, that the curve for $\delta=0.75$ lies parallel to the Morse curve and 0.05 above it at $R=1.4$, while the same is true of the curve for $\delta=0.875$ for $R=1.6$. In view of the small effect of changes in δ, it seems safe to conclude that, within the region studied, and to an accuracy of 0.01 v.e., the envelope of a family of curves having various δ values would be given by calculating the energy for each R with $2\delta=1.08R$, and that this envelope would lie 0.05 v.e. above the observed Morse curve. It follows that the calculated equilibrium distance and fundamental vibration frequency agree sensibly with the observed. It must be admitted, however, that the cogency of this reasoning would be much re-enforced by an actual computation with $\delta=0.625$, say, in order to exclude the conceivable possibility that for some R near 1.4 a still lower result might come out.

The coefficients of each term in the series were computed for all cases. They vary in a regular manner as R and δ are changed, showing no such sharp maxima or minima as the parameters in Rosen's paper. They are shown in Table V.

OTHER POSSIBLE APPLICATIONS

The excited Σ states of H_2 which, like the normal state, are singlet, with wave functions symmetrical in the nuclei, can be attacked by the same general method. Hylleraas and Undheim[9] have shown that the second lowest root of the secular equation is an upper limit for the energy of the second lowest state of the given symmetry, and so on. Thus, by investigating the values of the higher roots for various choices of R and δ we may hope to gain approximations to the potential curves for the excited states. Of course, this may demand the use of more terms than are needed for the ground state, or even the use of the exponential $e^{-\delta_1\lambda_1-\delta_2\lambda_2}$ in place of $e^{-\delta(\lambda_1+\lambda_2)}$. The various singlet Σ states antisymmetric in the nuclei can be investigated in the same way, by using those terms in the series with $j+k$ odd, and the triplet Σ states by using the series in which differences instead of sums occur:

$$\psi=\Sigma_{mnjkp}C_{mnjkp}e^{-\delta(\lambda_1+\lambda_2)}(\lambda_1{}^m\lambda_2{}^n\mu_1{}^j\mu_2{}^k\rho^p - \lambda_1{}^n\lambda_2{}^m\mu_1{}^k\mu_2{}^j\rho^p).$$

Once the computation for the singlet states has been carried out, the triplet states can be treated with very little labor. Some work on excited H_2 has been done in this laboratory, and will be communicated later. So far, the results are disappointing.

In applying a similar method to molecules with but two electrons outside closed shells, one would attempt to represent the inner shells by the usual atomic functions, the valence electrons by a series function. When only terms with $p=0$ are used, the work is straightforward and not too difficult. The introduction of the terms with r_{12} involves the evaluation of some exceedingly formidable integrals, and it seems that some modification of these terms will be needed. This is now being investigated in relation to the normal states of LiH and Li_2.

[9] E. A. Hylleraas and B. Undheim, Zeits. f. Physik **65**, 759 (1930).

Mathematical Appendix

Evaluation of matrix components

Consider the matrix components of unity and of energy arising between any two terms in the series from which the wave function is constructed, say the terms $[m_a n_a j_a k_a p_a]$ and $[m_b n_b j_b k_b p_b]$. Each term contains two sub-terms, which we may designate ψ_a, ψ_a', and ψ_b, ψ_b', respectively; they differ by interchange of indices between electrons. The matrix component of unity will have the form

$$S_{ab}=\int\int dV_1 dV_2 \psi_a\psi_b+\int\int dV_1 dV_2 \psi_a\psi_b'+\int\int dV_1 dV_2 \psi_a'\psi_b+\int\int dV_1 dV_2 \psi_a'\psi_b'.$$

It will be sufficient to consider the first of these integrals, which we may call s_{ab}. Four similarly related integrals will form the matrix component of the energy; the first of these will be

$$h_{ab}=\int\int dV_1 dV_2 \psi_a H\psi_b$$

$$=\int\int dV_1 dV_2 \psi_a[-\tfrac{1}{2}\nabla_1^2-\tfrac{1}{2}\nabla_2^2+(1/R)\{1+2/\rho-4\lambda_1/(\lambda_1^2-\mu_1^2)-4\lambda_2/(\lambda_2^2-\mu_2^2)\}]\psi_b.$$

Owing to the fact that the complete terms $\psi_a+\psi_a'$ and $\psi_b+\psi_b'$ contain the coordinates of the two electrons symmetrically, we may use instead the simpler integrals

$$h'_{ab}=\int\int dV_1 dV_2 \psi_a[-\nabla_2^2+(1/R)\{1+2/\rho-8\lambda_2/(\lambda_2^2-\mu_2^2)\}]\psi_b.$$

Then

$$H_{ab}=h'_{ab}+h'_{a'b'}+h'_{ab'}+h'_{a'b}.$$

In performing these integrations, we were not able to utilize the device adopted by Hylleraas for He, using ρ directly as one of the variables of integration. We found it necessary to transform the powers of ρ into elliptical coordinates, by means of the equations

$$\rho^2=\lambda_1^2+\lambda_2^2+\mu_1^2+\mu_2^2-2-2\lambda_1\lambda_2\mu_1\mu_2-2[(\lambda_1^2-1)(\lambda_2^2-1)(1-\mu_1^2)(1-\mu_2^2)]^{\frac{1}{2}}\cos(\varphi_1-\varphi_2),$$

$$\rho^{-1}=\sum_{\tau=0}^{\infty}\sum_{\nu=0}^{\infty} D_\tau^\nu P_\tau^\nu\binom{\lambda_1}{\lambda_2} Q_\tau^\nu\binom{\lambda_2}{\lambda_1} P_\tau^\nu(\mu_1) P_\tau^\nu(\mu_2)\cos\nu(\varphi_1-\varphi_2),\ \text{(notation of Zener and Guillemin[10])}$$

$$D_\tau^0=2\tau+1;\quad D_\tau^\nu=(-1)^\nu\times 2(2\tau+1)[(\tau-\nu)!/(\tau+\nu)!]^2 \text{ for } \nu>0.$$

Now, the required quantities may all be found in terms of integrals defined thus:

$$X(m,n,j,k,p)\equiv\frac{1}{4\pi^2}\int\int\int\int\int\int(\lambda_1^2-\mu_1^2)e^{-2\delta(\lambda_1+\lambda_2)}\lambda_1^m\lambda_2^n\mu_1^j\mu_2^k\rho^p d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\varphi_1 d\varphi_2.$$

The indices are all zero or positive, except that p may have the value -1.

Remembering that $dV_1 dV_2=(1/64)R^6(\lambda_1^2-\mu_1^2)(\lambda_2^2-\mu_2^2)d\lambda_1 d\lambda_2 d\mu_1 d\mu_2 d\varphi_1 d\varphi_2$, we find at once

$$s_{ab}=(1/64)R^6[X(m_a+m_b, n_a+n_b+2, j_a+j_b, k_a+k_b, p_a+p_b)$$
$$-X(m_a+m_b, n_a+n_b, j_a+j_b, k_a+k_b+2, p_a+p_b)].$$

We shall abbreviate this to

$$s_{ab}=(1/64)R^6[X(02000)-X(00020)].$$

---

[10] C. Zener and V. Guillemin, Jr., Phys. Rev. **34**, 999 (1929).

In h_{ab}', the only troublesome part is that arising from the Laplacian. Insofar as the coordinates of electron 2 are concerned, we may write ψ_b as constant $\times L_b(\lambda_2)M_b(\mu_2)P_b(\rho)$. Then

$$\nabla_2^2\psi_b/\psi_b=\nabla_2^2L_b/L_b+\nabla_2^2M_b/M_b+\nabla_2^2P_b/P_b+2\nabla_2L_b\cdot\nabla_2P_b/L_bP_b+2\nabla_2M_b\cdot\nabla_2P_b/M_bP_b;$$

(since the term in $\nabla_2L_b\cdot\nabla_2M_b$ always vanishes, the coordinates being orthogonal). Now, in general,

$$\nabla^2f(\rho)=[4/R^2\rho^2]\times[2\rho df/d\rho+\rho^2d^2f/d\rho^2];$$

$$\nabla^2f(\lambda,\mu)=[4/R^2(\lambda^2-\mu^2)]\times[2\lambda\,\partial f/\partial\lambda+(\lambda^2-1)\partial^2f/\partial\lambda^2-2\mu\,\partial f/\partial\mu+(1-\mu^2)\partial^2f/\partial\mu^2];$$

$$\nabla f(\lambda,\mu)\cdot\nabla g(\lambda,\mu,\varphi)=[4/R^2(\lambda^2-\mu^2)]\times[(\lambda^2-1)(\partial f/\partial\lambda)(\partial g/\partial\lambda)+(1-\mu^2)(\partial f/\partial\mu)(\partial g/\partial\mu)].$$

Putting

$$L_b=e^{-\delta\lambda_2}\lambda_2{}^{n_b},\qquad M_b=\mu_2{}^{k_b},\qquad P_b=\rho^{p_b},$$

we find

$$[(\lambda_2{}^2-1)/\lambda_2]\partial P_b/\partial\lambda_2=\tfrac{1}{2}p_bP_b\rho^{-2}[\rho^2-\lambda_1{}^2+\lambda_2{}^2-\mu_1{}^2-\mu_2{}^2+2\lambda_1\mu_1\mu_2/\lambda_2]$$

$$[(1-\mu_2{}^2)/\mu_2]\partial P_b/\partial\mu_2=-\tfrac{1}{2}p_bP_b\rho^{-2}[\rho^2-\lambda_1{}^2-\lambda_2{}^2-\mu_1{}^2+\mu_2{}^2+2\lambda_1\lambda_2\mu_1/\mu_2]$$

and finally

$$\psi_a\nabla_2^2\psi_b=[4\psi_a\psi_b/R^2(\lambda_2{}^2-\mu_2{}^2)]\times\{2(n_b-\delta\lambda_2)+(1-1/\lambda_2{}^2)(n_b(n_b-1)-2\delta n_b\lambda_2+\delta^2\lambda_2{}^2)$$
$$-2k_b+(1/\mu_2{}^2-1)(k_b(k_b-1))+(p_b/\rho^2)[(p_b+1)(\lambda_2{}^2-\mu_2{}^2)$$
$$+(n_b-\delta\lambda_2)(\rho^2-\lambda_1{}^2+\lambda_2{}^2-\mu_1{}^2-\mu_2{}^2+2\lambda_1\mu_1\mu_2/\lambda_2)-k_b(\rho^2-\lambda_1{}^2-\lambda_2{}^2-\mu_1{}^2+\mu_2{}^2+2\lambda_1\lambda_2\mu_1/\mu_2)]\}.$$

This can evidently be integrated in terms of the functions X previously defined. But a somewhat shorter expression for the integral can be found. By Greene's theorem, for properly continuous functions which vanish canonically at infinity,

$$\int dV_2\psi_a\nabla_2^2\psi_b=\tfrac{1}{4}\int dV_2(\psi_a\nabla_2^2\psi_b+\psi_b\nabla_2^2\psi_a-2\nabla_2\psi_a\cdot\nabla_2\psi_b).$$

We note

$$\nabla_2\psi_a\cdot\nabla_2\psi_b/\psi_a\psi_b=\nabla_2L_a\cdot\nabla_2L_b/L_aL_b+\nabla_2M_a\cdot\nabla_2M_b/M_aM_b+\nabla_2P_a\cdot\nabla_2P_b/P_aP_b+\nabla_2L_a\cdot\nabla_2P_b/L_aP_b$$
$$+\nabla_2M_a\cdot\nabla_2P_b/M_aP_b+\nabla_2L_b\cdot\nabla_2P_a/L_bP_a+\nabla_2M_b\cdot\nabla_2P_a/M_bP_a.$$

When this is evaluated, some cancellation occurs, and the final formula is

$$h'_{ab}=(1/64)R^5\{X(02000)-X(00020)+2X(0200-1)-2X(0002-1)-8X(01000)\}$$
$$-(1/64)R^4\{[(n_a-n_b)^2-(k_a-k_b)^2+(n_a+n_b)-(k_a+k_b)+(p_a-p_b)(n_a-n_b-k_a+k_b)]X(00000)$$
$$-4\delta X(01000)-[(n_a-n_b)^2-(n_a+n_b)]X(0-2000)+[(k_a-k_b)^2-(k_a+k_b)]X(000-20)$$
$$+[(p_a-p_b)^2+p_a+p_b+(p_a-p_b)(n_a-n_b+k_a-k_b)][X(0200-2)-X(0002-2)]$$
$$-(p_a-p_b)(n_a-n_b-(k_a-k_b))[X(2000-2)+X(0020-2)]$$
$$+2(p_a-p_b)(n_a-n_b)X(1-111-2)-2(p_a-p_b)(k_a-k_b)X(111-1-2)\}.$$

(The same abbreviation as before has been made in the argument of the X's.)

In discussing the computation of these integrals, it will be expedient to introduce the abbreviation $\alpha=2\delta$. Let us define the more general function

$$X^\nu(m,n,j,k,p)\equiv Z^\nu(m+2,n,j,k,p)-Z^\nu(m,n,j+2,k,p);$$

$$Z^\nu(m,n,j,k,p)\equiv(1/4\pi^2)\int\int\int\int\int\int e^{-\alpha(\lambda_1+\lambda_2)}\lambda_1{}^m\lambda_2{}^n\mu_1{}^j\mu_2{}^k\rho^pM^\nu\cos^\nu(\varphi_1-\varphi_2)d\lambda_1d\lambda_2d\mu_1d\mu_2d\varphi_1d\varphi_2;$$

$$M^2=(\lambda_1{}^2-1)(\lambda_2{}^2-1)(1-\mu_1{}^2)(1-\mu_2{}^2).$$

This evidently reduces to the previously defined $X(m, n, j, k, p)$ when $\nu=0$.
For $p=0$, we have

$$Z(m, n, j, k, 0)=4A_m(\alpha)A_n(\alpha)/(j+1)(k+1),$$

$$Z'(m, n, j, k, 0)=0,$$

$$Z''(m, n, j, k, 0)=8[A_{m+2}(\alpha)-A_m(\alpha)][A_{n+2}(\alpha)-A_n(\alpha)]/(j+1)(j+3)(k+1)(k+3),$$

when j and k are both even; otherwise the integrals vanish. (Here, as usual, $A_n(\alpha)\equiv\int_1^\infty \lambda^n e^{-\alpha\lambda}d\lambda$. For details of computation and tables, see Rosen's paper.[11])

For $p=-1$, the Neumann expansion of $1/\rho$ yields

$$Z(m, n, j, k, -1)=\sum_{\tau=0}^{\infty}(2\tau+1)R_\tau(j)R_\tau(k)H_\tau(m, n, \alpha),$$

$$Z'(m, n, j, k, -1)=-\sum_{\tau=1}^{\infty}[1/\tau(\tau+1)]R_\tau'(j)R_\tau'(k)\times[(2\tau+1)/\tau(\tau+1)]H_\tau'(m, n, \alpha),$$

$$\begin{aligned}Z''(m, n, j, k, -1)=&\tfrac{1}{2}\sum_{\tau=2}^{\infty}[1/(2\tau+3)(2\tau-1)(\tau+2)(\tau+1)\tau(\tau-1)]R_\tau''(j)R_\tau''(k)\\&\times[(2\tau+3)(2\tau+1)(2\tau-1)/(\tau+2)(\tau+1)\tau(\tau-1)]H_\tau''(m, n, \alpha)\\&+\tfrac{1}{2}\sum_{\tau=0}^{\infty}(2\tau+1)(R_\tau(j)-R_\tau(j+2))(R_\tau(k)-R_\tau(k+2))\\&\times[H_\tau(m+2, n+2, \alpha)-H_\tau(m+2, n, \alpha)-H_\tau(m, n+2, \alpha)+H_\tau(m, n, \alpha)].\end{aligned}$$

Here, $R_\tau^\nu(j)\equiv\int_{-1}^{+1}d\mu(1-\mu^2)^{\nu/2}P_\tau^\nu(\mu)\mu^j$. Rosen has discussed and tabulated this integral for $\nu=0$. The other cases are readily obtained from the elementary theory of associated harmonics. They exist only for even or for odd τ, according as $j+\nu$ is even or odd, and vanish for all values of τ greater than $j+\nu$, so that the summation over τ contains at most a few terms.

The other integral

$$H_\tau^\nu(m, n, \alpha)\equiv\int_1^\infty d\lambda_1\int_1^\infty d\lambda_2[(\lambda_1^2-1)(\lambda_2^2-1)]^{\nu/2}P_\tau^\nu\binom{\lambda_1}{\lambda_2}Q_\tau^\nu\binom{\lambda_2}{\lambda_1}\lambda_1^m\lambda_2^n e^{-\alpha(\lambda_1+\lambda_2)}$$

is familiar in molecular problems, at least for $\nu=0$. Rosen gives a method for computing $H_0(m, n, \alpha)$ (his formula A19). With the aid of his function $S(m, n, \alpha)$ (formulas A13, A14), and the following recurrence relations, other values can be found readily:

$$H_1(m, n, \alpha)=H_0(m+1, n+1, \alpha)-S(m, n+1, \alpha)-S(n, m+1, \alpha).$$

$$\begin{aligned}H_\tau(m, n, \alpha)=&(1/\tau^2)[(2\tau-1)^2H_{\tau-1}(m+1, n+1, \alpha)+(\tau-1)^2H_{\tau-2}(m, n, \alpha)\\&-(2\tau-1)(2\tau-3)(H_{\tau-2}(m+2, n, \alpha)+H_{\tau-2}(m, n+2, \alpha))+2(2\tau-1)(2\tau-5)H_{\tau-3}(m+1, n+1, \alpha)\\&-(2\tau-1)(2\tau-7)(H_{\tau-4}(m+2, n, \alpha)+H_{\tau-4}(m, n+2, \alpha))+2(2\tau-1)(2\tau-9)H_{\tau-5}(m+1, n+1, \alpha)\\&-\cdots]\text{ until either}\end{aligned}$$

$$\begin{aligned}&\text{(for even } \tau)\quad -(2\tau-1)[H_0(m+2, n, \alpha)+H_0(m, n+2, \alpha)-S(m+1, n, \alpha)-S(n+1, m, \alpha)]\\&\text{(for odd } \tau)\quad +(2\tau-1)[2H_0(m+1, n+1, \alpha)-S(m, n+1, \alpha)-S(n, m+1, \alpha)].\end{aligned}\qquad\text{for } \tau>1.$$

$$[(2\tau+1)/\tau(\tau+1)]H_\tau'(m, n, \alpha)=(\tau+1)H_{\tau+1}(m, n, \alpha)-(2\tau+1)H_\tau(m+1, n+1, \alpha)+\tau H_{\tau-1}(m, n, \alpha).$$

[11] N. Rosen, Phys. Rev. **38**, 255 (1931).

$$[(2\tau+3)(2\tau+1)(2\tau-1)/(\tau+2)(\tau+1)\tau(\tau-1)]H_\tau''(m,n,\alpha)$$
$$=\tau(2\tau-1)[(2\tau+3)/(\tau+1)(\tau+2)]H'_{\tau+1}(m,n,\alpha)$$
$$-(2\tau+3)(2\tau-1)[(2\tau+1)/\tau(\tau+1)]H_\tau'(m+1,n+1,\alpha)+(\tau+1)(2\tau+3)[(2\tau-1)/(\tau-1)\tau]H'_{\tau-1}(m,n,\alpha).$$

(The factors in brackets [ ] before H' and H'' are left in explicitly because for purposes of computation they may conveniently be considered part of the functions.)

When X^ν has been tabulated for $p=0$ and $p=-1$, it can easily be found for positive values of p by means of the relation (with the usual abbreviation)

$$X^\nu(00002)=X^\nu(20000)+X^\nu(02000)+X^\nu(00200)+X^\nu(00020)-2X^\nu(00000)$$
$$-2X^\nu(11110)-2X^{\nu+1}(00000).$$

A corresponding relation holds also for Z^ν.

Solution of secular equation, and determination of constants

The complexity of the secular equation of course necessitates determination of the roots by trial and error. The following process seems to be the simplest available, and puts the work in a particularly satisfactory form for the determination of the constants. Choosing a trial value of λ in the proper neighborhood (known from experiment, or experience with a similar function), we may write the s equations which must be satisfied by the coefficients C in the form

$$\sum_{\tau=1}^{s} d_{i\tau}C_\tau=0, \qquad i=1,\,2,\,\cdots\,s, \qquad d_{i\tau}=H_{i\tau}-\lambda S_{i\tau}.$$

This set of equations is equivalent to another set, obtained from it by progressive elimination of the variables C_τ, in which there is one equation involving s of the C's, one in $(s-1)C$'s, and so on. Let these equations be

$$\sum_{\tau=1}^{s} D_{i\tau}C_\tau=0, \qquad i=1,\,2,\,\cdots\,s.$$

The D's are conveniently found by the recurrence relation

$$D_{n\tau}=d_{n\tau}-\sum_{\nu=1}^{n-1} D_{\nu\tau}D_{\nu n}/D_{\nu\nu},$$

starting with $D_{1\tau}=d_{1\tau}$. Because the d's form a symmetrical matrix, it will be found that $D_{n\tau}$ vanishes automatically for $\tau<n$. A convenient numerical check is provided by the fact that the quantities $\Sigma_\tau D_{n\tau}$ should be obtainable from the quantities $\Sigma_\tau d_{n\tau}$ by means of the same recurrence relation.

If λ has been properly chosen, the last factor D_{ss} will vanish, and the last equation $D_{ss}C_s=0$ will permit C_s to be set arbitrarily equal to 1 (or any other number). Since the secular determinant is easily seen to be just the product of the diagonal factors D_{ii}, the value of λ so determined is evidently that which makes the determinant vanish. The remaining C's may be immediately determined by successively substituting C_s in the $(s-1)$st equation, C_s and C_{s-1} in the $(s-2)$nd, and so on. The provisional coefficients are then used to compute $\int\int dV_1 dV_2\psi^2$; upon division by the square root of this quantity, a set of normalized coefficients is obtained.

Should D_{nn} vanish for $n<s$, the meaning is that the value of λ chosen happens to be an exact solution of the secular equation containing only the first n terms. In this case (and also in practise if D_{nn} is very small), the procedure for subsequent D's breaks down. If, nevertheless, it is desired to find the value of the whole determinant, the order of the original terms must be changed, placing at least one of the originally first n terms in a position later than the nth.

In general, of course, the chosen value of λ will prove incorrect, the last factor, and therefore the whole determinant, failing to vanish. In order to find a value of λ making D_{ss} negligibly small, the above process is carried out for two assumed values of λ lying rather near the root, and preferably enclosing it, as shown by the last factor having opposite signs in the two cases. The root can then be quite accurately determined by considering the value of the determinant (not the last factor) as linear in λ, and interpolating. If the coefficients are desired, they must be obtained by a new computation using the interpolated λ, as it is not safe to try to interpolate the individual D's which determine the coefficients.

It may be mentioned that by this method an 11-row determinant with six significant figures can be evaluated and checked by an experienced computer in a little over two hours.

Order of magnitude of $(H-E)\psi$

Suppose our approximate wave function ψ expanded in terms of the correct wave functions ψ_n of proper symmetry, with coefficients a_n, such that a_1 is nearly unity, and $\Sigma_n a_n^2=1$. Let E be the energy calculated, and E_n the correct energy for the nth function. Then

$$E=\int \psi H\psi dV=\int\left(\sum_\tau a_\tau\psi_\tau\right)\left(\sum_\sigma a_\sigma E_\sigma\psi_\sigma\right)dV=\sum_\tau a_\tau^2E_\tau=a_1^2E_1+\sum_{\tau=2} a_\tau^2E_\tau$$

$$=a_1^2E_1+\bar{E}\sum_{\tau=2}a_\tau^2=a_1^2E_1+\bar{E}(1-a_1^2),$$

where $\bar{E}$ is some value between E_2 and E_∞. Similarly we can show that

$$\delta^2=\int[(H-E)\psi]^2dV=a_1^2(E_1-E)^2+(\bar{\bar{E}}-E)^2(1-a_1^2),\qquad E_2<\bar{\bar{E}}<E_\infty.$$

Eliminating $1-a_1^2$ gives $\delta^2=(E_1-E)[E_1-E+\{(\bar{\bar{E}}-E)^2-(E_1-E)^2\}/(E_1-\bar{E})]$.

Assuming that $(\bar{\bar{E}}-E)/(\bar{E}-E_1)$ is of the order of unity, and neglecting higher powers of $\epsilon=E-E_1$, we have $\delta^2\approx(E_1-E)(E_1-\bar{E})$. This is a lower limit, since it can easily be shown that $\bar{\bar{E}}>\bar{E}$.

Note added in proof:

A minute error has been found, affecting the energy by 0.003 e.v. Corrected coefficients will be published as soon as possible.

Dielectric Constant Studies. I. An Improved Voltage Tuning Resonance Method and Its Application to Aqueous Potassium Chloride Solutions

J. GILBERT MALONE,* A. L. FERGUSON AND L. O. CASE, *Chemistry Laboratory, University of Michigan*

(Received August 30, 1933)

An experimental study of the behavior of various types of apparatus used in the measurement of dielectric constants of aqueous solutions of electrolytes. Although simple theory indicates that the voltage tuning method of detection should give results which are uninfluenced by the conductivity of the solutions and should, hence, be particularly adapted to measurements on electrolytic solutions, it appears, however, that other disturbing factors are present. This is shown by the divergent results obtained when different types of detection are employed and leads to the conclusion that the change in dielectric constant of such solutions as observed by this method depends more or less upon the characteristics of the particular generating and receiving circuits used.

THE experimental evidence for the effect of ions on the dielectric constant of liquids is very conflicting. Debye,[1] after a consideration of the results, obtained prior to 1929, states that, "no definite conclusion with regard to any theory can be drawn from measurements in their present condition." Since 1928, numerous papers have appeared and one might expect considerable improvement; however, the pronounced lack of agreement in published data forces the conclusion that this is not the case. The seriousness of this situation was emphasized to the authors in the course of some work on the dielectric constant of gelatin solutions, the results of which will be reported in a later paper. In connection with that work, it became evident that in some solutions conductivity had an important bearing on the results. These circumstances led the authors to undertake a series of investigations in an attempt to clarify the situation in respect to aqueous solutions of electrolytes. This aim has necessitated the use of several methods of measurement in order to contrast directly their advantages and limitations.

It is the purpose of this paper to discuss the so-called voltage tuning resonance method used in various forms and with varying success by Jezewski,[2] Kniepkamp,[3] Lattey and co-workers,[4] and Astin.[5] Since the presentation of this work in the form of a dissertation, another paper by Jezewski[6] has appeared which describes work similar in its purpose to that presented here.

Since the theory has been given by Jezewski, and can also be found in the papers of Lattey already referred to, it need be described but briefly here. The voltage tuning principle derives its significance from the fact that analysis of the tunable portion of the circuit of Fig. 1 shows that, if the value of C is adjusted to give a maximum voltage across the condenser, this maximum will be independent of the resistance x in parallel

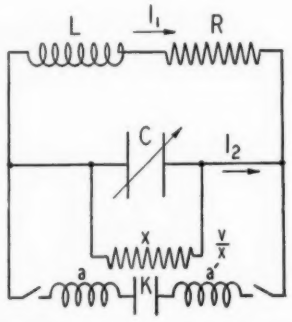

FIG. 1. Diagrammatic resonating circuit.

* This paper constitutes a portion of the dissertation submitted by J. Gilbert Malone to the faculty of the Graduate School, University of Michigan, in partial fulfillment of the requirements for the degree of Sc.D.

[1] Debye, *Polar Molecules*, p. 124, Chem. Cat. Co., N. Y., 1929.

[2] Jezewski, Zeits. f. Physik **48**, 123 (1928).

[3] Kniepkamp, Zeits. f. Physik **51**, 95 (1928).

[4] Lattey (*et. al.*), Phil. Mag. **7**, 985 (1929); **11**, 997 (1931); **12**, 1111 (1931); **13**, 444 (1932).

[5] Astin, Phys. Rev. **34**, 300 (1929).

[6] Jezewski, Phys. Zeits. **34**, 88 (1933).

with it. This analysis assumes an inductance without capacitance, and a condenser without inductance. Moreover, there must be no reaction of the resonance circuit on the oscillator.

The Oscillator

The oscillations were produced by the use of a UX210 tube in the circuit shown in Fig. 2.

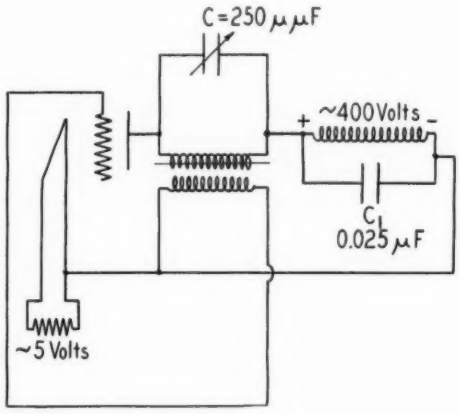

Fig. 2. Oscillator.

The filament was lighted and the plate supplied with the proper voltages by a transformer unit especially designed for this purpose. The inductances of the plate circuit and the grid circuit were wound on the same Bakelite core and fitted with plug connections facilitating their interchange. UX bases are convenient for this purpose as they will not permit the coil so mounted to be placed wrongly in the circuit. This oscillator functioned well for wave-lengths from 20 to 100 meters.

The Resonating Circuit

When the voltage tuning resonance method is employed there are several factors to be considered which do not appear from the simple theory. The simple theory takes no account of the inductance of the cell leads. It is obvious, however, from consideration of the circuit shown in Fig. 1 that the inductance will be changed because of the inductance in the leads represented by a and a' if the total capacity is transferred from the cell K to the precision condenser C. It is evident also that some sort of a meter must be used to indicate the maximum voltage across the condenser C, and whatever sort is used, it presents complications by introducing capacitance, inductance, resistance or some combination of these. Lattey and his co-workers employ a thermocouple galvanometer connected across the condenser terminals through a capacitive impedance. They believe this arrangement to be superior to all others for the purpose.

Two distinct resonating circuits have been used in this work. The first followed that of Lattey and Davies[7] as closely as possible, and is shown in Fig. 3. This circuit as constructed and

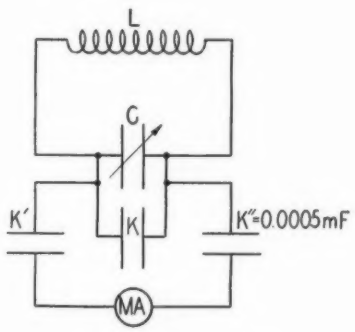

Fig. 3. First receiver.

used by the present authors, gave results in qualitative agreement with those obtained by Lattey and Davies. Since the second circuit is believed to represent an advance, and since it possesses far greater sensitivity and does not sacrifice or compromise any principle embodied in the first, no further discussion of the first circuit is considered necessary.

The second circuit is shown in Fig. 4. In principle it is an alternating voltage indicating device. In this arrangement it is not necessary to balance out an initial plate current as is usually done; the plate current is initially zero because of the d.c. grid bias. When the resonance condition is approached an increase in the plate current of the tube results and can be measured by a milliammeter in the plate circuit. When the voltage across the condenser is a maximum, the plate current is likewise a maximum, and both of these conditions occur, as has been pointed out by Jezewski, when $\omega^2 = 1/LC$.

It is desirable that the parts of the resonator

[7] Lattey and Davies, Phil. Mag. **12**, 1111 (1931).

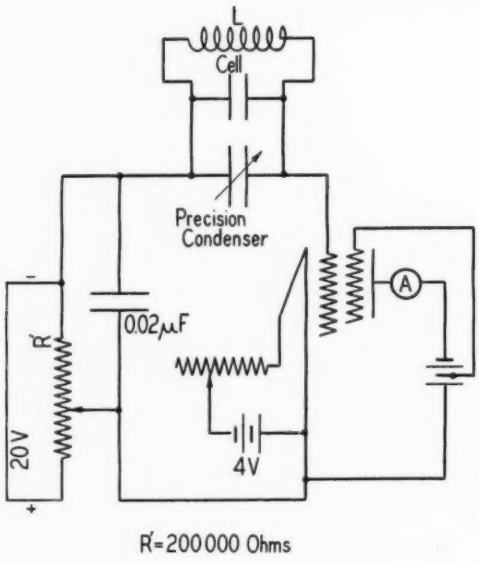

FIG. 4. Second receiver.

be arranged as compactly and as symmetrically as possible. The inductances were made from 8×2 mm copper strip and were mounted on short extensions of the condenser terminals. The tube connections were placed at about the same point, and provision was made to allow the cell to be inserted directly across the terminals of the condenser and the tube connections. It was possible to bring all of these connections together by inclining the condenser to the extent that a radio tube could be hung just below its terminals. The extension of these terminals provided a means for securing the inductance and the cell.

The condenser used was General Radio type 222-L. A UX222 tube was employed. All leads and the tube were shielded and grounded. The dielectric cells used are illustrated by the one shown in Fig. 5. These were made of glass and platinum with very short leads. It has been concluded by Lattey that in such a case the apparent dielectric capacity will not be affected by the conductivity of the solution.

The oscillator was normally kept at a distance of about 90 cm from the resonator. Tests showed that effects of coupling were observable at 30 cm. It has been observed also that unless one works at a considerable distance from conductors such as water pipes, etc., effects due to these may be attributed to coupling. On the basis of such observations the authors believe that the changing of the relative positions of oscillator and resonator is undesirable.

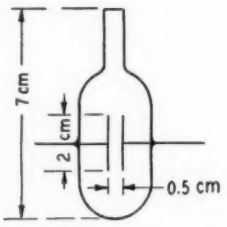

FIG. 5. Dielectric constant cell.

The setting of the condenser, giving maximum current, which is the condition of resonance, is difficult to determine with great precision and is often obtained, in the case of wave meters, by the insertion of a small capacitance in parallel with the variable condenser. Any setting of the condenser such that insertion of this capacitance makes no change in the reading of the indicating device requires that the capacitance has spanned the resonance point to give a corresponding point on the other side. This procedure, while admirable for wave meters, was not very successful here and an alternative was sought.

It is generally known that the curve obtained when current is plotted against condenser readings is symmetrical if the calibration curve of the condenser is linear. To what extent this is true in the present case may be seen from the figures in Table I.

TABLE I. *Data on symmetry of resonance curve.*

Microamperes	Condenser Values Above	Below	Mean
400	5390	5260	5325
300	5431	5230	5330
200	5470	5172	5329
150	5520	5122	5321
100	5744	4864	5304
	Cell filled with solution		
110	5104	5502	5303
140	5278	5322	5300

The first column gives the current as shown by the ammeter *A* in microamperes, the next two columns give the two condenser settings on opposite sides of the resonance point corresponding to the currents in the first column. The last column shows the mean of these condenser values. If the above noted symmetrical relation-

ship holds, the values in the last column should remain constant. The extent of the agreement indicates that the symmetrical relationship is closely approximated until the curve is about 400 condenser divisions in width. The values obtained with an electrolyte show that the curve is flattened as a result of electrolytic conduction. Only a limited range may, therefore, be followed and the accuracy is correspondingly decreased.

If the calibration curve of the condenser were linear, a number of readings might be made at corresponding current values on either side of the resonance point, and the resonance point obtained as above by taking the mean. In practice it was found necessary to use the lower settings of the condenser where the linear relationship does not hold and in such cases three procedures are possible. In the first method, the condenser may be calibrated and the readings converted in terms of the calibration. This method was not employed in the present case, since experience indicated two other methods which were more convenient and of equal accuracy. In both the average of the two condenser readings obtained with the same current but on opposite branches of the resonance curve was plotted against the difference between these readings, i.e., against the width of the resonance curve. The two methods differed, however, in the fact that in method II the plate current was always adjusted to the same value while in method III it was allowed to vary.

The procedure used in taking measurements according to method II was as follows: Two condenser readings were taken, one on each side of the resonance point for the same plate current. The condenser was then set to a new value and, after the plate current had been restored to the original value by suitable variation of the grid bias, another condenser reading was taken on the opposite side of the resonance point. This procedure was then repeated for as many points as desired.

In method III, on the other hand, the plate current was allowed to vary with the closeness of approach to the resonance point, although again for each PAIR of readings the plate current was of course the same.

These two methods gave satisfactory agreement over the node width of 200 to 600 divisions, which range included practically all of the measurements in this work. In practice, however, method III was preferred in that it gave somewhat more consistent results.

In applying the latter method to the measurement of the dielectric constant of an electrolytic solution, curves were determined for water and for the solution in the manner just described and both plotted on the same graph as illustrated in Fig. 6. The curves obtained approximate straight

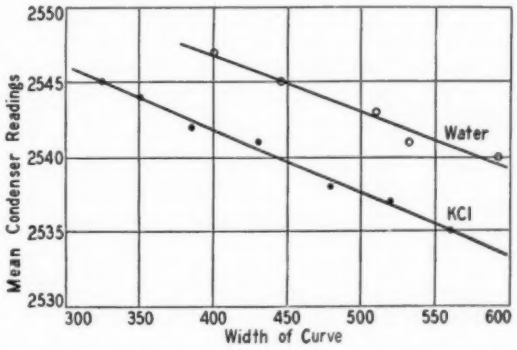

FIG. 6. Plots of typical observational series for the dielectric constant of KCl solutions by the voltage tuning method.

lines and the curve for the solution always falls below the one for water. The mean ordinate distance between these curves is the increase in the capacity of the cell when filled with solution over its capacity when filled with water. The dielectric constant of the solution is calculated from this value. Following this method of procedure, the troublesome conversion of condenser readings into comparable units is avoided and at the same time the effect of an asymmetric wave is minimized.

The apparatus and cells described were designed for the measurement of the dielectric constants of aqueous solutions of electrolytes. It was assumed that the relationship between cell capacity and dielectric constant is linear, and that benzene and water can be employed as calibrating materials to determine this line. Several pure materials, covering a wide range of dielectric constants, were measured. The values obtained agreed well with those recorded in the literature.

The Dielectric Constant of Aqueous Solutions of KCl

A review of this subject has been given by Lattey and Davies.[7] Many observers have measured a large number of solutions.[8] The conflicting nature of the results obtained emphasizes the futility of making extensive measurements on solutions of a large number of carefully purified inorganic substances at the present time. It has been the purpose of the present authors to confine measurements to a few substances and to study the possibilities of several methods. The Drude method in several modified forms has been used by us, and is the subject of a separate communication.*

Lattey and his co-workers have used the voltage tuning resonance method most extensively. Some of his methods, on first consideration, appear different from and more elegant than those employed by us. Careful analysis, however, reveals that this is probably not the case. Let us consider the problem of finding the true resonance point. The impressed voltage under proper conditions, as has been indicated, gives a potential across the condenser which is symmetrical with respect to capacity and falls off with capacity values either less or greater than this. In order to obtain K_0, the resonance capacity, Lattey considers that two values for the capacity are obtained, one on either side of the resonance capacity K_0, which give the corresponding currents I_1 and I_1' and a current I_0 at the resonance point. He is thus enabled to write:

$$(K_0-K_1)/(I_0^2/I_1^2-1)^{\frac{1}{2}}=\phi$$

and

$$(K_1'-K_0)/(I_0^2/I_1'^2-1)^{\frac{1}{2}}=\phi.$$

Eliminating K_0:

$$\phi=(K_1'-K_1)/(I_0^2/I_1^2-1)^{\frac{1}{2}}+(I_0^2/I_1'^2-1)^{\frac{1}{2}}$$

ϕ is the tuning factor.

Lattey and Davies determine current values with the condenser set at definite values and calculate K_0 by means of the preceding equations. It is obvious from the derivation of the equation that, if $I_1'=I_1$, the values of K_0 will be the arithmetical mean of K_1 and K_1'. The method followed in our work consists in reading the values of capacitance, first with the current at a definite value and then after the current has been brought to the same value on the other side of the resonance point by a proper adjustment of the capacity. Such a procedure is more convenient in observation and calculation than that used by Lattey and Davies and is more fundamentally sound in that the accuracy of the indicating device does not introduce a source of error. This is of special significance when a thermionic tube is employed as the voltmeter to measure the potential across the condenser.

The measurements of Lattey and Davies show that the precision was not high. Thus, with a solution of approximately 0.001 N, they show a variation of from 1.5 to 3.0 percent and with a 0.12 N solution the variation is about 10 percent in the extreme cases. Nevertheless, the trend of the values reported by Lattey is definite; and the first consideration in this work is to obtain results which, even though qualitative in nature, are fundamentally sound.

Astin's results[5] with KCl by the resonance method are represented by three curves, the first showing a decrease to a minimum at a concentration of about 0.001 N followed by a rapid increase to values above that of water, while the other two, with closer coupling, show a continued decrease, more rapid with a cylindrical plate condenser than with a parallel plate condenser.

The data obtained by the present authors do not confirm those of Astin but show instead a continued rise in the dielectric constant with concentration. The observed data are recorded in Table II. These values were taken at 25°C at a wave-length of 31.7 meters. Cell A had a capacity of 606 $\mu\mu$f and cell B a capacity of 28.2 $\mu\mu$f when filled with water.

The value for pure water has been accepted as the reference value. Only low concentrations could be measured with the larger cell due to conductivity. The apparent dielectric constant change due to solute is small yet definite in direction, fairly reproducible and increases with increasing concentration.

A concentration of 0.0025 N solution gave

[8] Compare for instance the results of Milicka and Slama, Ann. d. Physik (5) **8**, 663 (1931) with those of Lattey and Davies, Phil. Mag. **12**, 1111 (1931); **13**, 444 (1932).

* This communication is really number II in this series and follows the present article.

TABLE II.* *Summary of dielectric constants (uncorrected) of aqueous KCl solutions.*

Conc.	Cell *A*		Cell *B*			
	1	2	1	2	3	4
0.0000	78.58	78.58	78.58	78.58	78.58	78.58
0.0010	78.13	79.13	79.45	79.42	79.42	79.17
0.0020		79.69	79.93	80.95	80.27	80.44
0.0025		81.21	80.27		81.97	81.12
0.0050			81.63			85.71

* *Note:* Each value in each column is the average of series of measurements made by refilling the cell with the same solution. The results in the different columns were made with freshly prepared solutions, at different times and under somewhat different experimental conditions, in order to determine whether there were any uncontrolled factors which might be having an effect.

dielectric constants of 81.57, 81.35 and 81.42 at wave-lengths of 24.5, 46.5 and 82.4 meters, respectively. These values seem to indicate that the results are independent of frequency.

The values here recorded for KCl are in agreement essentially with those obtained by Lattey and Davies. They are in contradiction to the values generaliy obtained with the Drude method.

It has been pointed out that the theoretical treatment assumes a pure inductance. Every conductor, however, has a certain amount of capacitance associated with it. The authors attempted to change the distributed capacitance effect by immersing the inductance coil of the measuring apparatus in benzene. Results obtained on salt solutions with the coil immersed gave, within the experimental error, no change of dielectric constant with concentration of the salt. The dielectric constant of all solutions, within the experimental error, under these conditions, was that of pure water. The benzene was not replaced with materials of higher dielectric constant for the reason that conductivity would be a factor in the use of materials such as the alcohols and the amount of material required and the experimental conditions made the use of volatile or expensive materials impractical. Even if the curve of apparent dielectric constants against the dielectric constant of the milieu surrounding the inductance is established the extrapolation to zero dielectric constant, i.e., to zero distributed capacity, is probably not justifiable. It was not practical to surround all of the inductance with benzene, and the problem may be one of reflection effects at the surface of the benzene. This circumstance, while it renders the problem even more complicated, does not change the basic fact that the distributed capacity of the coil is a factor of importance in using the voltage tuning resonance method and, as yet, has not been satisfactorily taken into account.

The results given are consistent with the values obtained by Jezewski as well as qualitatively in agreement with those of Wien.[9] In contradiction to Jezewski the present authors are convinced, however, that, even though their data recorded in Table II indicate a fairly definite behavior for aqueous KCl solutions, yet, when all things are considered, the voltage tuning resonance method cannot be depended upon to distinguish between the effects of increased conductivity and dielectric constant change in aqueous salt solutions.

[9] Wien, Ann. d. Physik **11**, 429 (1931).

Dielectric Constant Studies. II. The Drude Method Applied to Aqueous Solutions of Potassium Chloride

J. GILBERT MALONE,* L. O. CASE AND A. L. FERGUSON, *Chemistry Laboratory, University of Michigan*

(Received August 30, 1933)

This paper is an experimental comparison of several adaptations of the Drude method as applied to conducting solutions. With three distinct types of detecting device, results for the dielectric constant of water and of water-alcohol mixtures agreed within experimental error. When, however, the same apparatus was applied to measurements with KCl solutions, different results were obtained with each. This indicates that any apparent change in the dielectric constant of electrolytic solutions, as observed by the Drude method, must depend in whole or in part upon the characteristics of the apparatus.

WITH the intention of measuring the dielectric constants of some electrolytic solutions, the authors constructed several modifications of the Drude apparatus. Comparison of the results obtained with these modifications, by using the same electrolyte, indicated a pronounced ambiguity, which, judging from our own experience and from work reported in the literature, seems to be characteristic, also, of the results of other methods.

Of the methods which have been employed for the measurement of dielectric constant of leaky dielectrics, the Drude method has much in its favor. It has, however, three obvious limitations: first, it requires a large amount of material; second, the material used must be in the liquid state; and third, because of the length of the tube required it can be operated only at very high frequencies. On the other hand, it appears to possess the best theoretical foundation, since the dielectric constant of a leaky dielectric is defined directly by the Maxwell equation:

$$(\epsilon\mu/c^2)\partial^2E/\partial t^2+(4\pi\rho\mu/c^2)\partial E/\partial t=\nabla^2E,$$

where ρ is the conductivity of the medium, ϵ its dielectric constant, μ its magnetic permeability, c the velocity of light and E the electric intensity. The method has been applied to electrolytic solutions by several investigators, among whom may be mentioned Drude,[1] Deubner,[2] Devoto,[3] Drake, Pierce and Dow.[4] In general the results with this method were interpreted as indicating practically no change of dielectric constant with concentration of the electrolyte. (An exception might be noted in the result of Deubner[2] for $CuSO_4$; he found a decrease in the dielectric constant with an increase in concentration, a result in disagreement with the earlier observations of Drude on the same material.) This, it will be noted, is at variance with the results of other methods, some of which[5] seem to indicate an increase and others[6] a decrease in the dielectric constant with concentration of electrolyte.

The essentials of the Drude apparatus are a source of high frequency oscillations, a Lecher wire system and a means of detecting the nodal or antinodal points on the wires.

The high frequency oscillations used by Drude[1] and other earlier workers consisted of damped waves. Deubner,[2] and Drake, Pierce and Dow[4] employed oscillations from the usual re-

* Rewritten from a thesis presented by J. Gilbert Malone in partial fulfillment of the requirements for the degree of Doctor of Science at the University of Michigan.

[1] Drude, Zeits. f. physik. Chemie **23**, 267 (1897).
[2] Deubner, Ann. d. Physik **84**, 429 (1927).
[3] Devoto, Gazz. chim. ital. **60**, 208 (1930).
[4] Drake, Pierce and Dow, Phys. Rev. **35**, 613 (1930).
[5] Wien, Ann. d. Physik (5) **11**, 429 (1931). Lattey and co-workers, Phil. Mag. **7**, 985 (1929); **11**, 997 (1931); **12**, 1111 (1931); **13**, 441 (1932). Jezewski, Phys. Zeits. **34**, 88 (1933).
[6] Walden, Ulich and Werner, Zeits. f. physik. Chemie **129**, 389 (1927). Shanke and Schreiner, Phys. Zeits. **28**, 597 (1927). Sack, Phys. Zeits. **28**, 199 (1927). Bramley, J. Frank. Inst. **205**, 649 (1928).

generative thermionic oscillators, while Devoto[3] used oscillations from a Barkhausen-Kurtz generator.

The earlier form of the Lecher wire system used by Drude and by Deubner consisted of the conventional arrangement of two parallel wires surrounded by the medium to be investigated. Devoto[3] modified the Drude method by the use of a single wire instead of two. Drake, Pierce and Dow[4] substituted a concentric cylinder system for the parallel wires.

The detecting device used by Drude was a neon bulb, Devoto and Deubner employed a milliammeter of the thermocouple type, while Drake, Pierce and Dow used the reaction of the Lecher wire system on the oscillator to give the nodal points.

The present authors have employed two types of Lecher wire systems, three distinct types of detection and a source of undamped waves.

Experimental

A diagram of the apparatus is shown in Fig. 1. For measurements of the wave-length in air, two wires, (see Fig. 1a) W and W', 1.5 inches apart, were mounted vertically between the floor and ceiling of the room, the ends being secured by Bakelite supports. The lower ends were connected to the terminals of a Weston milliammeter while a shorting bridge B, constructed of metal and Bakelite and provided with spring contacts, was movable on the wire system. The oscillator was placed on the floor within a foot or two of the wire system, and the positions of the nodal points of the waves induced on the wires were determined by moving the bridge and noting the points of maximum current as shown by the milliammeter.

For the measurement of the wave-length in liquid media a similar arrangement (Fig. 1c) was employed except that in this case the wires were $\frac{1}{2}$ inch apart and were surrounded for the greater part of their length by a glass tube 5 inches in diameter and about 7 feet long. The lower end of the tube was closed by a rubber stopper, through which the wires protruded to the milliammeter. The connections of the wires at the upper end of the tube are shown in the figure. In order to maintain a constant temperature, the liquid was

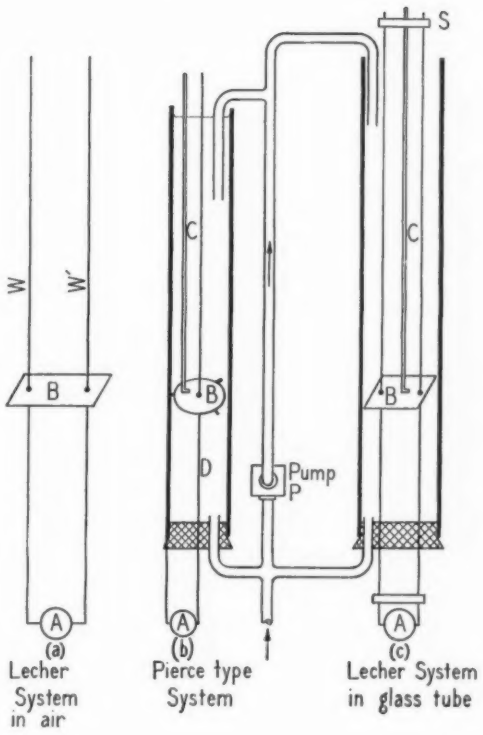

Fig. 1. Diagram of apparatus.

circulated by means of an electrically driven pump connected with glass and rubber tubing to the top and bottom of the large glass tube. Heating units were located around the tube and the whole was heat insulated. A thermometer in one of the glass connections from the pump gave a convenient means of reading the temperature. The shorting bridge B providing reflection of the wave within the tube was moved from the top by means of a calibrated Bakelite rod (C), which, by means of a vernier, was read to a precision of 0.01 cm. The ammeter at the bottom was read from the top by means of a telescope.

In addition to the glass tube apparatus just described the system of Drake, Pierce and Dow was also employed. This apparatus (Fig. 1b) consisted of a brass tube 7.5 feet long and 2.5 inches in diameter, closed at the lower end by a rubber stopper and containing an axially located phosphorbronze wire D. The circulation of the liquid and the detecting arrangement used here were of the same type as described above for the glass tube.

All wires, the interior of the metal tube, and the exposed parts of the pump were heavily silver plated, and where parts were exposed to considerable wear solid silver was used. In the first detecting arrangement employed in the present work, the ammeter was fixed at the generator end of the wire system and the shorting bridge moved to locate the maxima. This arrangement has been used by Laville.[7]

The undamped waves employed were generated by a vacuum tube oscillator. The measurements were carried out in the following manner. The wave-length in air was first determined with the auxiliary wire system. In the early part of the work these measurements were checked on each of the other systems, but it soon became evident that this was not necessary. Without changing the generator the wave-length in the liquid was then determined with either the glass or brass tube arrangement. The wave-length in air was remeasured at the end of each experiment to make sure that it had not changed. The ratio of the half wave-lengths thus obtained was taken as the index of refraction, n. For nonconductive media $n^2 = \epsilon$.

With the object of comparing the two tubes and of testing the accuracy of the apparatus, some preliminary measurements were made with water and with water-alcohol mixtures. The results with water are shown in Table I. It will be noted that the glass tube gave slightly lower values for ϵ than the concentric cylinder system. This result was attributed to the fact that a sensible portion of the field lay outside the water in the tube. Support was given to this explanation by the still lower value of 72.90 obtained when the Lecher wires were further apart.

The results obtained with alcohol-water mixtures covering the entire concentration range agreed satisfactorily with those of Wyman[8] the average deviation being 0.75 percent.

It seemed that the above determinations established the reliability of the apparatus for work with non-electrolytes and accordingly some work on electrolytes was attempted. For these measurements, KCl and $CuSO_4$, representing two types of ions, were selected, solutions of these salts having been frequently investigated by others. The results obtained are given in Fig. 2. The values given are for n^2; since in this work we are concerned only with relative values, no attempt has been made to apply a correction for conductivity, which would be uncertain at best. It will be noted that the values show a decided decrease with increase in concentration, whereas others, working with the Drude method, as was pointed out earlier, have found practically no effect on the dielectric constant due to electro-

TABLE I. *Comparison of parallel wire and concentric tube systems.*

Temp.	λ/2 air	λ/2 in liquid	ε at T	ε at 25°C
20.4	176.2	*g* 19.74	79.68	78.03
		m 19.67	81.25	78.59
16.6	176.2	*g* 19.55	81.23	78.61
water changed		*m* 19.54	81.31	78.69
21.2	176.2	*g* 19.74	79.68	78.36
		m 19.74	79.68	78.36
24.2	176.2	*g* 19.86	78.71	78.36
		m 19.85	78.79	78.54
29.0	176.2	*g* 20.06	77.15	78.59
water changed		*m* 20.04	77.31	78.75
23.3	176.1	*g* 19.85	78.70	78.09
		m 19.77	79.34	78.73
24.0	176.2	*g* 19.92	78.24	77.88
		m 19.87	78.64	78.28
23.2	176.2	*g* 19.86	78.71	78.11
		m 19.81	79.11	78.51
30.4	176.2	*g* 20.15	76.46	78.40
		m 20.13	76.61	78.55
water changed				
20.1	176.2	*g* 19.69	80.08	78.22
		m 19.67	80.24	78.48
22.6	176.2	*g* 19.78	79.35	78.49
		m 19.77	79.43	78.57
26.5	176.2	*g* 19.96	77.93	78.47
		m 19.95	78.00	78.68
water changed				
23.6	176.2	*g* 19.86	78.71	78.20
		m 19.80	79.19	78.69
26.0	176.2	*g* 19.94	78.01	78.43
		m 19.93	78.16	78.52
25.0	176.2	*g* 19.92	78.24	78.24
		m 19.86	78.71	78.71

Average as measured by glass tube 78.36.
Average as measured by metal tube 78.58 at 24°C.
m is value measured by metal tube system.
g refers to the glass container and the Lecher wire system.

[7] Laville, *Les Ondes Electriques Courtes*, p. 126, R. Mesny, Paris (1927).

[8] Wyman, J. Am. Chem. Soc. **53**, 3292 (1931).

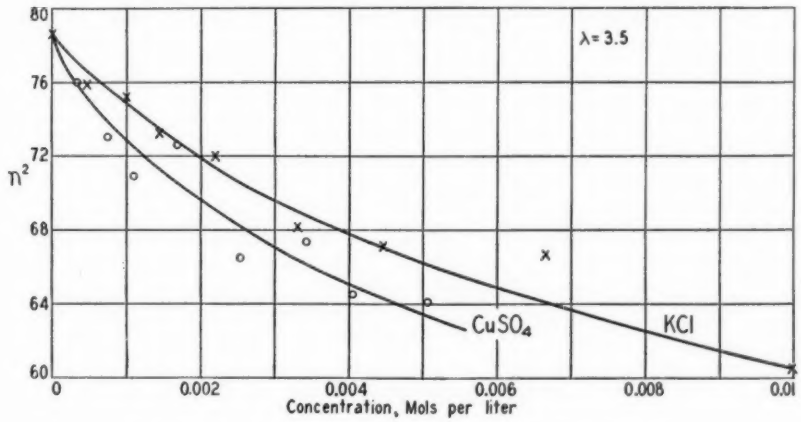

FIG. 2. n^2 against concentration for electrolytic solutions.

lytes. These discrepancies may be due at least in part to differences in the equipment used. For instance, the various detecting devices no doubt introduce complicating factors in the application of simple theory, factors which may produce results differing both in magnitude and direction from the correct n^2 value. No quantitative theoretical relationship taking these factors into account has so far been developed.

An experimental study of this question was undertaken by the use of two additional types of detecting device. The first was a voltmeter constructed from a UX222 tube and which had been found satisfactory in connection with some resonance work being carried on in this laboratory. The potentials across the input end of the apparatus could be followed by the changes in the plate current of the tube. The grid was connected to the center wire, forming the inner cylinder of the concentric cylinder system, while the filament was connected to the tube by a short heavy wire and was also grounded. This apparatus functioned very well but it was more sensitive to changes in the surroundings than was the case with the previous detector. With this device, the values obtained for water and for one concentration of ethyl alcohol, 34.5 percent, were consistent with those previously obtained. While in these cases the position of the nodes was altered by the substitution of the new detector, the distance between nodes was the same as before. However, the precision was much less than in the case already described. With electrolytes, however, greatly different results were obtained. Although the results were not very reproducible, a large and fairly uniform *increase* of n^2 with increasing concentration was observed.

A third detecting device was used, depending upon the reaction of the measuring system upon an additional resonant circuit, the changes in which were observed by means of the vacuum tube voltmeter which was inserted across the input terminals. With this arrangement the nodes in the measuring circuit could be observed by noting the points of maximum reaction on the added circuit as shown by the plate current of the vacuum tube in it. Here again the values observed for alcohol and water were consistent with those obtained in the preceding cases, but the values observed for KCl solutions were quite different. The n^2 values calculated from the observations show again an *increase* with increasing concentration of electrolyte, but less than in the preceding case. The limited extent of this increase indicates that there is little change in dielectric constant when it is calculated from n^2 by means of the Maxwell relationship for a leaky dielectric. Such a result approximates that obtained by Drake, Pierce and Dow using a similar method.

The last two procedures were not carried to high refinement by the present authors and although they gave fairly consistent results they were particularly sensitive to changes of capacitance in the surroundings.

It is generally realized that the problem of the dielectric constant of electrolytic solutions is highly complex. This was the conclusion arrived

at by Debye[9] in 1929 as a result of a critical study of the literature up to that time. As stated in the first paper of this series, the authors undertook this work in an attempt to clarify this situation. The results show, however, that the problem is even more complex than has been realized. The measurements reported in this and the preceding paper indicate that while self-consistent results can be obtained even for electrolytes by each of the various methods in the hands of different observers, there is no conformity in the results of different methods, even when employed by the same investigator.[10] Since, as has been pointed out, these discrepancies do not appear in the case of non-electrolytic solutions, it is evident that conductivity is responsible for the anomalous results obtained, and that its effect is different with different methods. The few attempts which have been made to treat this effect theoretically are inadequate, and until an adequate treatment has been made the changes observed by any method cannot be regarded as representing true changes in dielectric constant.

[9] Debye, *Polar Molecules*, Chapter VI, Chemical Catalog Co., New York (1929).

[10] cf. Lattey, Phil. Mag. **41**, 829 (1921); **12**, 1111 (1931); **13**, 444 (1932).

The Heat of Formation of Binary Liquid Solutions from Their Liquid Components

J. J. COLEMAN AND FRANK E. E. GERMANN, *Department of Chemistry, University of Colorado*

(Received October 9, 1933)

It is pointed out that an equation giving the relation of the heat of formation per mol of liquid solutions from their liquid components, to temperature and composition can be used as the foundation equation for their thermodynamical theory. The empirical equation

$$\Delta H = N_1(1-N_1)(\alpha+\beta e^{\gamma N_1}),$$

in which ΔH is the heat of formation per mol of the solution, N_1 is the mol fraction of one of the components of the binary solution, and α, β and γ are constants or functions of temperature, is presented. There are two degenerate forms of this equation. (Eqs. (7) and (8).) There are listed values of α, β and γ (or the values of the corresponding quantities in Eqs. (7) and (8)) for 59 solutions for one temperature and sometimes for several temperatures. It is noted that according to the data assembled in the *International Critical Tables*, some solutions composed of "unassociated" liquids present a behavior more irregular than those composed of "associated" liquids. It is suggested that this unexpected behavior may mean that some of these data are not as precise as they are indicated to be.

I. Use of the Heat of Formation Equation as the Foundation Equation for the Complete Thermodynamical Theory of the Solution

IT has long been known that many of the problems of solutions can be reduced to that of finding the dependence of their thermodynamical properties upon temperature and composition. This has been done for dilute solutions, for "ideal" solutions,[1] for van der Waals solutions[2] (i.e., those in which the van der Waals equation can be used as the equation of state of the components of the solutions in both the liquid and gaseous state) and for one or two other types of solutions. Only the theory of van der Waals solutions has any general application and this yields only qualitative results.

We cannot get all of these relations from thermodynamics alone. We must have at least one of them to start with. In the case of "ideal" solutions we can take as the fundamental extra-thermodynamical relation

$$f_1 = N_1 f_1^0 \qquad (1)$$

where f_1 is the fugacity of any component in the solution, N_1 is its mol fraction, and f_1^0 is its fugacity in a pure phase of the same kind (solid, liquid or gaseous) as that of the solution.

Now by the methods of K. Watanabe[3] we can derive the equation, valid for any component in any solution

$$\ln f_1 = \ln N_1 f_1^0 - \int_{(P,\ N_1)} \frac{\overline{H}_1 - H_1^0}{RT^2} dT, \qquad (2)$$

in which $\overline{H}_1$ is the partial molal heat content of the component in the solution, H_1^0 is its molal heat content in the pure phase, T is the temperature and R is the gas constant. The integration is indicated to be at constant pressure and composition. This integral is evidently the correction term necessary to make (1) (written in the logarithmic form) applicable to non-ideal solutions. And the $\overline{H}_1 - H_1^0$ appearing in the integrand can be related to the heat of formation of the solution by differentiation of the obvious equation

$$H_s - n_1 H_1^0 - n_2 H_2^0 = \Delta H(n_1+n_2)$$

(in which H_s is the heat content of a solution containing n_1 mols of component one and n_2 mols of component two and ΔH is the heat of formation *per mol* of the solution) with respect to

[1] Lewis and Randall, *Thermodynamics*, Chapters XIX and XX.

[2] van der Waals and Kohnstamm, *Lehrbuch der Thermostatik*, Part II.

[3] Watanabe, *On the Vapour Pressure of Liquid*, Tokyo Science Reports, Sect. A, No. 6 (1931).

n_1(T, p and n_2 constant). This gives

$$\begin{aligned}\bar{H}_1-H_1^0 &= (n_1+n_2)\left[\frac{\partial \Delta H}{\partial n_1}\right]_{n_2,\ T,\ p}+\Delta H \\ &= (n_1+n_2)\left[\frac{\partial \Delta H}{\partial N_1}\right]_{n_2,\ T,\ p}\left[\frac{\partial N_1}{\partial n_1}\right]_{n_2,\ T,\ p}+\Delta H \qquad (3) \\ &= (1-N_1)\left[\frac{\partial \Delta H}{\partial N_1}\right]_{T,\ p}+\Delta H.\end{aligned}$$

(There is a similar equation for solutions of more than two components but we are concerned with binary solutions only in this article.)

It is evident then that Eq. (2) enables us to use any relation giving the dependence of ΔH on N_1 and T as the fundamental extra-thermodynamical equation for non-ideal solutions. But before turning to such equations we will give a derivation of (2).

For the change

dn_1 (pure liquid)

$+(n_1+n_2)$ (solution)$\rightarrow(n_1+n_2+dn_1)$ (solution)

we have a change of free energy and of heat content

$$dn_1(\bar{F}_1-F_1^0), \qquad dn_1(\bar{H}_1-H_1^0).$$

These two quantities are related by an equation from the general theory of thermodynamics

$$\left[\partial\frac{(\bar{F}_1-F_1^0)dn_1}{T}\Big/\partial T\right]_{N_1,\ p}=-\frac{(\bar{H}_1-H_1^0)dn_1}{T^2}.$$

Here dn_1, the amount dissolved, is, of course, independent of T. Cancelling this out and integrating we get (after dividing by R)

$$\begin{aligned}\frac{\bar{F}_1-F_1^0}{RT} &= \ln\frac{f_1}{f_1^0} \\ &= -\int_{(N_1,\ p)}\frac{\bar{H}_1-H_1^0}{RT^2}dT+g(N_1, p), \qquad (4)\end{aligned}$$

in which $g(N_1, p)$ is an arbitrary function.

Consider now the solution under any pressure p and any temperature T. If we first increase the pressure until it exceeds the critical pressure of the solution and then increase the temperature until it exceeds the critical temperature, we will have passed by a series of *continuous* states from a liquid solution to a gaseous solution. It seems safe to assume that the form of the function $g(N_1, p)$ is not altered in the first step of this process and that it is not changed at all by the second step; and since the gaseous solution is ideal or can be made so by a further increase in temperature we can use the relations for ideal solutions

$$\bar{H}_1-H_1^0=0, \quad f_1=N_1f_1^0,$$

which give us

$$g=\ln N_1.$$

Substitution into (4) then gives the desired Eq. (2).

II. The Baud-Heitler Equation

In 1915 Baud[4] brought forth the equation

$$\Delta H=bN_1(1-N_1), \qquad (5)$$

in which b is a constant, and showed that his equation served for a number of solutions. Later Heitler[5] gave a derivation of this equation based upon statistical mechanics and gave some of the thermodynamical relations for solutions for which this equation holds. Hildebrand[6] also used this equation in developing the thermodynamics of a new class of solutions which he has called "regular solutions." It is probably the most serviceable equation for the heat of formation of binary solutions that has been previously developed.

III. A New Heat of Formation Equation

It is evident then that the study of the function $\Delta H(T, N_1)$ offers a promising approach to the problem of extending the thermodynamical theory of solutions. We have sought first to find an equation for the relation between ΔH and N_1 which would serve for solutions in which the

[4] Baud, *Analyse Thermique des Melanges Binaires*, Bull. Soc. Chim. (4) **17**, 329–45 (1915).

[5] Heitler, Ann. d. Physik (4) **80**, 630 (1926).

[6] Hildebrand, J. Am. Chem. Soc. **51**, 66 (1929).

Baud-Heitler one fails. A purely empirical analysis of the results assembled in the *International Critical Tables* resulted in the following equation:

$$\Delta H = N_1(1-N_1)(\alpha+\beta e^{\gamma N_1}), \qquad (6)$$

where α, β and γ are either constants or functions of temperature. This formula was found to have a wide application, but before turning to this we will consider some interesting special cases.

If α and β are of opposite sign then it may happen that ΔH changes in sign as we proceed over the composition range. This is a frequent phenomenon. If γ is small enough to enable us to use the approximation $e^x = 1+x$ then Eq. (6) reduces to

$$\Delta H = N_1(1-N_1)(\alpha_1+\beta_1 N_1), \qquad (7)$$

in which α_1 and β_1 are new constants or functions of temperature. If either β or γ is zero then the equation reduces to the Baud-Heitler form

$$\Delta H = N_1(1-N_1)\alpha_2. \qquad (8)$$

Finally if α, β and γ are all zero we have, of course, the case of the "ideal" solution.

In Table I we have listed values for α, β and γ (or α_1 and β_1, or α_2 as the case may be) for a number of solutions. In the first four columns we give these and the temperature (or interval within which the temperature lies). In columns five to eight we give some data which are intended to give some notion of how well the equation serves. In column five is listed the maximum value of a deviation index which is the numerical value of

$$\frac{\Delta H \text{ given by the equation} - \Delta H \text{ given by experiment}}{\text{Average of experimental values (all considered positive)}} \times 100.$$

We did not use the percentage deviation for if we had sought to make this a minimum, it would have meant giving the greater weight to observations of smaller numerical value, where the relative precision is lower. We assumed that, unless otherwise indicated, the absolute precision was the same for all measurements. Hence we sought to keep the absolute deviations as small as possible and our index indicates how well we were able to do this. In some cases all of the data except one could be represented nicely by our equation. In these cases we ignored this one result in the determination of α, β and γ and the index of deviation. Hence we have given in columns six and seven the number of observations ignored and the total number of them. (In this last we included the two zero values of ΔH at $N_1=1$ and $N_1=0$.) In some cases while what data there were could be easily represented by our equation, their distribution over the composition range was so poor as to make uncertain the applicability of the equation or, assuming this, the values for α, β and γ given. Hence in column eight we have commented on the distribution of the data.

In Figs. 1 and 2 we give two illustrations. In Fig. 1 we have plotted the heat of formation of CS_2 - $C_{10}H_{16}$ (pinene) solutions against the mol fraction of CS_2. The solid line is the plot of the

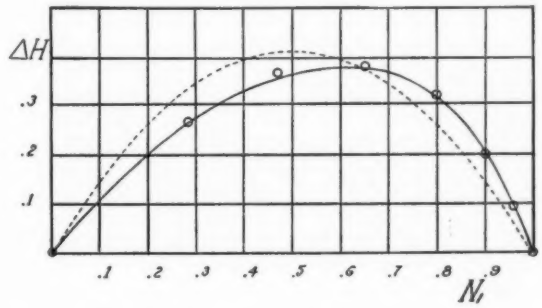

FIG. 1. Variation of heat of formation of $CS_2-C_{10}H_{16}$ solutions with mol fraction of CS_2.

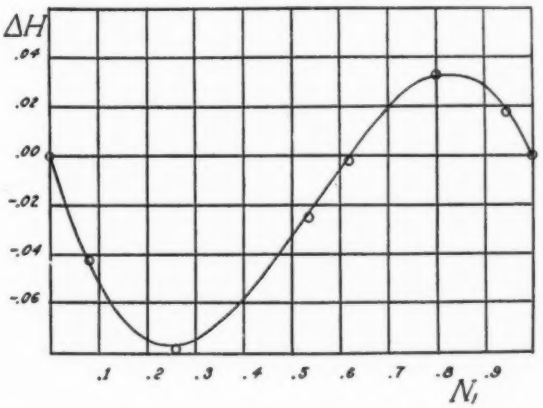

FIG. 2. Variation of heat of formation of $C_2H_4Cl_2-C_7H_8$ solutions with mol fraction of $C_2H_4Cl_2$.

TABLE I.

(The heat unit used is the kilojoule.)

Component two	t°C	α, α_1 or α_2	β or β_1	γ	Maximum index of deviation	Number of observations ignored	Total number of observations	Distribution	Observer
H_2O (Component one)									
CH_4O Methyl alcohol	0	− 2.95	− .0525	5.00	3	0	21	Good	1
CH_4O Methyl alcohol	19.69	− 2.55	− .103	4.00	9	0	21	Good	1
CH_4O Methyl alcohol	42.37	− 1.33	− .234	3.00	3	0	21	Good	1
C_2H_6O Ethyl alcohol	0	− 1.75	− .038	6.06	5	0	21	Good	1
$C_2H_6O_2$ Glycol	32	− 1.40	− .232	3.00	7	1	11	Good	2
$C_2H_6O_2$ Glycol	55	− 1.27	− .213	3.00	2	1	11	Good	2
$C_2H_6O_2$ Glycol	76	− 1.27	− .191	3.00	9	0	11	Good	2
CCl_4 (Component one)									
$C_2H_4Br_2$ Ethylene Bromide	25	1.92	.31	—	5	0	9	Good	3
C_6H_6 Benzene	25	.44	—	—	7	0	8	Good	4
C_6H_7N Aniline	25	4.67	.01	6.20	3	0	13	Good	4
C_7H_8 Toluene	17	.106	.106	—	7	0	11	Good	6
CS_2 (Component one)									
CCl_4 Carbon tetrachloride	25	1.14	.0239	3.00	2	0	11	Good	3
$CHCl_3$ Chloroform	25	1.93	.48	—	2	0	11	Good	3
$C_2H_4Br_2$ Ethylene bromide	25	2.52	.222	2.00	1	0	8	Good	3
$C_4H_8O_2$ Ethyl acetate	25	3.95	.159	2.70	1	0	9	Good	3
$C_4H_{10}O$ Ethyl ether	25	− 1.52	− .63	—	3	0	9	Good	3
C_6H_6 Benzene	4	1.77	.0531	3.00	3	0	11	Good	7
C_6H_6 Benzene	14.5	1.73	.055	3.00	2	0	10	Good	6
C_6H_6 Benzene	25	1.93	.67	—	4	0	8	Good	3
$C_6H_{12}O_2$ Paraldehyde	25	3.90	.222	3.00	7	0	9	Good	3
$C_{10}H_{16}$ Pinene	25	1.10	.0790	3.00	5	0	9	Good	3
$CHCl_3$ (Component one)									
CCl_4 Carbon tetrachloride	25	.890	.11	—	5	0	8	Good	3
C_2H_6O Ethyl alcohol	25	− 6.30	11.3	—	8	1	11	Good	3
$C_4H_8O_2$ Ethyl acetate	25	− 5.77	−4.55	—	5	0	10	Fair	3
$C_4H_{10}O$ Ethyl ether	25	−10.5	—	—	4	0	10	Good	3
$C_4H_{10}O$ Ethyl ether	14	− 9.96	—	—	7	1	11	Good	6
C_6H_5Cl Chlorobenzene	15–20	− .598	− .210	—	1	0	6	Poor	5
C_6H_6 Benzene	15–20	− 1.00	− .75	—	5	0	8	Good	5
C_6H_6 Benzene	25	− 1.40	− .70	—	8	0	8	Good	3
C_6H_{10} Cyclohexene	15–20	− .443	—	—	3	0	5	Poor	5
C_6H_{12} Cyclohexane	15–20	2.63	—	—	3	0	7	Fair	5
$C_6H_{12}O_2$ Paraldehyde	25	− 6.05	−4.65	—	5	0	9	Good	3
C_8H_{10} p-Xylene	15–20	− 3.23	− .80	—	1	0	8	Good	5
CH_4O Methyl alcohol (Component one)									
C_2H_6O Ethyl alcohol	3	.053	—	—	2	0	5	Fair	8
C_3H_8O n-Propyl alcohol	25	3.00	0.146	3.00	3	0	5	Fair	3
$C_4H_{10}O$ Ethyl ether	25	.57	2.48	—	4	0	9	Good	3
$C_5H_{12}O$ Isoamyl alcohol	25	.22	1.08	—	2	0	5	Fair	3
$C_2H_4Br_2$ Ethylene bromide (Component one)									
$C_2H_4O_2$ Acetic acid	?	2.20	.0722	3.00	2	0	10	Good	12
C_6H_6 Benzene	15–20	1.16	—	—	3	0	9	Good	5
C_6H_{12} Cyclohexane	15–20	5.56	—	—	3	0	14	Good	5
$C_2H_4Cl_2$ Ethylene chloride (Component one)									
C_7H_8 Toluene	15–20	− .690	1.11	—	4	1	9	Good	5
C_2H_4O Acetaldehyde (Component one)									
$C_4H_{10}O$ Ethyl ether	25	−2.29	—	—	2	0	8	Good	3
$C_2H_4O_2$ Acetic acid (Component one)									
C_6H_{12} Cyclohexane	15–20	6.77	—	—	4	1	9	Good	5
C_7H_8 Toluene	15–20	1.49	—	—	3	0	8	Fair	5
H_2O Water	?	2.25	−.376	2.00	9	0	11	Fair	9
C_2H_6O Ethyl alcohol (Component one)									
C_3H_6O Acetone	25	4.63	—	—	9	0	10	Good	3
C_3H_6O Acetone (Component one)									
CH_4O Methyl alcohol	25	2.32	.82	—	2	0	10	Good	3
$C_4H_8O_2$ Ethyl acetate	25	.544	—	—	4	0	9	Good	3
$C_3H_6O_2$ Propionic acid (Component one)									
H_2O Water	8	1.20	7.90	—	4	0	13	Good	9
$C_3H_6O_2$ Methyl acetate (Component one)									
$C_4H_8O_2$ Ethyl acetate	16	.279	—	—	3	1	11	Good	6
$C_4H_8O_2$ Ethyl acetate (Component one)									
C_2H_6O Ethyl alcohol	25	4.70	.114	3.00	4	0	9	Fair	3
C_3H_8O n-Propyl alcohol	25	5.88	.0106	6.00	3	1	10	Good	3
$C_4H_{10}O$ Ethyl ether	25	1.04	—	—	2	0	10	Good	3
$C_4H_{10}O$ Isobutyl alcohol	25	7.43	—	—	4	0	9	Good	3
$C_4H_{10}O$ Ethyl ether (Component one)									
CH_4O Methyl alcohol	25	.57	2.48	—	4	0	9	Fair	3
C_3H_8O n-Propyl alcohol	25	1.94	.189	3.00	4	0	10	Good	3
$C_4H_{10}O$ Isobutyl alcohol	25	2.27	.417	2.00	1	0	10	Good	3
$C_5H_{12}O$ Isoamyl alcohol	25	1.95	.397	2.00	6	0	10	Good	3
$C_6H_{12}O_2$ Paraldehyde	25	1.03	.45	—	5	0	9	Good	3
C_6H_5Cl Chlorobenzene (Component one)									
C_8H_{10} p-Xylene	15–20	− .48	—	—	5	0	7	Good	5
C_6H_6 (Component one)									
CH_4O Methyl alcohol	15	1.10	.240	3.00	10	0	11	Good	6
$C_2H_4O_2$ Acetic acid	?	1.73	.47	—	9	0	15	Good	12
C_6H_{12} Cyclohexane	15–20	3.31	—	—	8	0	7	Good	5
C_6H_{14} n-Hexane	15–20	3.86	—	—	4	0	6	Fair	5
C_7H_8 Toluene	15–20	.298	—	—	7	0	8	Good	5
C_7H_8 Toluene	16	.298	—	—	10	0	11	Good	6
C_6H_{11} Cyclohexane (Component one)									
C_7H_8 Toluene	15–20	2.43	.47	—	10	0	9	Good	5
C_8H_{10} p-Xylene (Component one)									
C_6H_7N Aniline	0	−3.80	−.150	[3.00]	3	0	8	Poor	11

OBSERVERS

[1] Bose, Zeits. f. physik. Chemie **58**, 585 (1907).
[2] Schwers, Rec. trav. Chim. **28**, 42 (1909).
[3] Hirobe, J. Faculty of Science (Tokyo) **1**, 155 (1926).
[4] Hartung, Trans. Faraday Soc. **12**, 66 (1917).
[5] Baud, Bull. Soc. Chim. de France **17**, 329 (1915).
[6] Schmidt, Zeits. f. physik. Chemie **121**, 211 (1926).
[7] Winkelmann, Ann. d. Physik **30**, 529 (1873).
[8] Bose, Ann. de Chim. et de Phys. **27**, 89 (1912).
[9] Faucon, Ann. de Chim. et de Phys. **19**, 70 (1910).
[10] Drucker and Moles, Zeits. f. physik. Chemie **75**, 405 (1911).
[11] Clark, Phys. Zeits. **6**, 154 (1905).
[12] Baud, Ann. de Chim. et de Phys. **27**, 89 (1912).

equation

$$\Delta H = N_1(1-N_1)(1.10+0.0790e^{3N_1}),$$

the dotted line is the plot of the equation

$$\Delta H = N_1(1-N_1)1.64$$

and the circles represent the experimental data. The dotted line indicates that in this case a Baud-Heitler equation will serve as a rough approximation. The maximum value of the deviation index (for the first equation) ignoring none of the experimental results is five. In Fig. 2 we give a similar graph for $C_2H_4Cl_2-C_7H_8$ (toluene) solutions. The equation is

$$\Delta H = N_1(1-N_1)(-0.690+1.11N_1).$$

The maximum deviation index, ignoring one of the experimental results, is four. The result ignored ($\Delta H = 0.661$, $N_1 = 0.385$) cannot be shown on the graph. The Baud-Heitler equation cannot, of course, be applied to solutions of this type.

There are a large number of cases in which the equation cannot be made to represent the data and it is quite evident that it is not perfectly general. However, solutions composed of such notoriously bad behavers as water and the alcohols gave points on a plot of $\Delta H/N_1(1-N_1)$ against N_1 which could be represented by a simple smooth curve and so it was expected that solutions of simpler (i.e., "unassociated") liquids would do the same. But these in many cases gave very irregular plots. This unexpected behavior suggests that it might be well to check some of these data.

There are not sufficient data available to enable us to attack the remaining part of the problem, the determination of the dependence of α, β and γ on T, in this direct empirical fashion without some preliminary experimental work. We could proceed at once to deduce thermodynamical relations for solutions for which Eq. (6) serves, leaving the α, β and γ in as undetermined functions of temperature. But the resulting equations are so clumsy that their application would be extremely difficult. Hence it seems best to delay this until we have further information concerning α, β and γ. One looks to statistical mechanics for an interpretation of this equation and such investigations when successful will, no doubt, throw some light on these quantities.

Contact Potentials and the Effects of Unimolecular Films on Surface Potentials. I. Films of Acids and Alcohols*

WILLIAM D. HARKINS AND E. K. FISCHER,† *Department of Chemistry, University of Chicago*

(Received August 19, 1933)

Simultaneous measurements of film pressure and surface potential have been carried out by the use of an apparatus designed in such a way that the potential may be determined for any location on the surface of the film. At film pressures above that of the gaseous films, organic substances with homo-heteropolar molecules give a single smooth curve for the relation between surface potential and molecular area. At areas sufficiently great to reduce the pressure to that of the gaseous film the surface potential becomes variable and remains variable until the area becomes so great that the continents and islands of condensed film evaporate in the two-dimensional system to give a gaseous film alone. For example with films of myristic acid at 17° the surface potential is represented by a single curve below a molecular area of about 50 sq. Å, and by any value below 170 mv at higher areas, at which islands in the film persist. The areas above which the surface potentials become variable, due to the effects of islands or continents of film, are found to be about as follows: stearic acid, 28; palmitic acid, 28; pentadecylic acid, 41; lauric acid (on NaCl-HCl solution), 36; oleic acid, 52; cetyl alcohol, 23.7; myristyl alcohol, 27. The surface potential rises rapidly with the number of carbon atoms in the hydrocarbon chain, and is about 400 mv for condensed films of stearic acid. It is higher for alcohols than for acids with the same number of carbon atoms. The increment is about that given by the addition of two carbon atoms to the acid. The change of one of the single bonds of stearic acid to the double bond of oleic acid as the film forming substance lowers the surface potential by from 37 to 65 percent, and makes this potential much more greatly affected by the molecular area. The following numbers give the number of carbon atoms in several normal acids and the corresponding potentials at 20 sq. Å at 17°: (12), 276; (14), 336; (15), 375; (16) 383; and (18), 410. The surface potential increases with decrease of temperature, with an increase of concentration in the film, and for the substances investigated, by a replacement of a basic by an acid solution. A collapse of the film lowers the surface potential since it changes the molecular orientation in the film. The films exhibit in the condensed state considerable hysteresis, both with respect to film pressure and surface potential. This feature has not been sufficiently noted in other work. The relation between surface potential and the molecular moments of the molecules which constitute the film is discussed. The dipole moment as calculated from surface potentials by the equation of Helmholtz are only about one-sixth as large as those found by ordinary methods for independent molecules in nonpolar solvents. The discrepancy is attributed partly to a lowering of potential by an orientation of the dipoles with respect to each other, and partly to the too great simplicity of the equation. The apparatus used for this work has been found to give good determinations of the contact potentials of metals. In its design use was made of ideas suggested principally by Lord Kelvin, and by Guyot and by Frumkin.

I. INTRODUCTION

IT has been supposed that the surface of a pure liquid or of a solution is the seat of several different types of potentials. Of these the greatest prominence in the theory of the stability of colloids, as developed by Hardy and later by others, has been given to the electrokinetic or zeta (ζ) potential, the original theory of which was developed by Helmholtz. It has not been found possible to measure the quantity designated by ζ directly as a potential, and it is necessary to introduce relative motion of the phases along the interface if an electromotive force (E) is to be developed, and this appears as a potential difference along and not through the interface. For this reason ζ is sometimes considered as a tangential potential, although it is E, which may have an entirely different order of magnitude, which is measured tangentially.

What may be considered as the contact potential of the surface is less fictitious, since it may be determined by a direct measurement. If the surfaces of two metals are separated by a gas, and if the gas is sufficiently ionized, a potential, known as the contact potential (V) is developed between

* The work reported in this article was carried out with the aid of a fellowship from the Julius Stieglitz Fund for Research in Chemistry Applied to Medicine, established at the University of Chicago by the Chemical Foundation. The objective of the present work is a study of membrane formation.

† Julius Stieglitz Fellow, University of Chicago.

the two metals. This particular method for the determination of contact potentials was developed by Kelvin.[1]

What may be designated as the surface potential (ΔV) of a solution was determined in 1883 by Bichat and Blondlot[2] and later by other workers.[3] The method used by Kenrick was to allow the reference liquid to flow down over the inner wall of a moderately large glass tube placed with its axis vertical, and to allow the second liquid to flow from a fine glass tip in a narrow jet down this axis. The two solutions were connected through calomel electrodes to a quadrant electrometer. It is claimed by both Kenrick and Frumkin that this method gives satisfactory results for aqueous solutions of inorganic electrolytes and for not too dilute solutions of derivatives of the hydrocarbons. The errors inherent in the method of the "dropping electrode" are discussed by Williams and Vigfusson.[4]

Unfortunately the values of the surface potentials as determined for similar systems by different observers did not exhibit a good agreement. In an attempt to obtain better values Guyot and Frumkin independently[5] developed a method in which the potential between a metal, coated with a radioactive substance, and a solution, is determined. The method of Guyot and Frumkin was used by Frumkin and Williams[6] and by Schulman and Rideal.[7]

The development of this method did not, however, remove the difficulty that the values obtained by different workers, or even by the same worker, on seemingly identical systems, were extremely variable. The researches on surface films which had been carried on for many years in this laboratory had shown that many films of insoluble organic substances of the polar-nonpolar type are much less homogeneous than had usually been supposed in the literature of the subject. From this point of view it was suggested to Frumkin by one of us (Harkins) in 1928 that the discrepancies would disappear if simultaneous measurements of potential and film pressure were to be taken, the latter by the use of a film balance of the type used by Adam. This suggestion, repeated by Frumkin to Rideal, was the origin of changes in methods abroad, and was taken as the basis of work in the present paper, which agrees remarkably well with that published recently by Adam and Harding.[8] The agreement is as good at low values for the area of the organic molecule as can be expected, especially when the great difficulty of obtaining really pure specimens of normal long chain acids is taken into account. At higher areas all of the difficulties disappear as will be shown later, if the potential is determined for all of the different regions of the film.

Although the relation between the Kelvin-Righi method and that of Guyot and Frumkin seems to have been obscured in the literature of the subject[9] it is obvious that the latter method is the same as the former, with the substitution in the latter of the surface of a solution for that of one of the metals in the former. Thus the potential determined by the method of Guyot and Frumkin is a contact potential. For obvious reasons in the determination of the effect of surface films it is more satisfactory to take the potential of the clean surface as an arbitrary zero, and to consider the change of potential caused by the presence of a film as the surface potential.

A new method for the determination of the contact potential of metals and of the related surface potential of a solution has been devised by W. A. Zisman,[10, 11, 12] who causes the metal elec-

[1] Sir Wm. Thomson, Lord Kelvin, *Mathematical and Physical Papers*, Vol. VI, *Contact Electricity of Metals*, pp. 110 to 145, Cambridge Press, 1911. See also Righi, J. de Physique [2] **7**, 153 (1888).

[2] Bichat and Blondlot, J. de Physique [2] **11**, 548 (1883).

[3] Kenrick, Zeits. f. physik. Chemie **19**, 625 (1896); Guyot, Ann. de Physique (10), **2**, 506 (1924); Frumkin, Zeits. f. physik. Chemie **111**, 190 (1924); **116**, 485 (1925); Garrison, J. Phys. Chem. **29**, 1517 (1925); Bühl, Ann. d. Physik [4] **81**, 211 (1927); [4] **87**, 877 (1928).

[4] Williams and Vigfusson, J. Phys. Chem. **35**, 348 (1931).

[5] Guyot, Ann. de Physique [10] **2**, 506 (1924). Frumkin, Zeits. f. physik. Chemie **116**, 485 (1925).

[6] Frumkin and Williams, Proc. Nat. Acad. Sci. **5**, 400 (1929).

[7] Schulman and Rideal, Proc. Roy. Soc. **A130**, 259, 270, 284 (1930).

[8] Adam and Harding, Proc. Roy. Soc. **A138**, 411 (1932). See also Schulman and Hughes, ibid. 430 (1932), and Fosbinder and Lessig, J. Frank. Inst. **215**, 425 (1933).

[9] See, however, a paper which has appeared since this paper was first written; Whalley and Rideal, Proc. Roy. Soc. **A140**, 484, 489, 497 (1933).

[10] Zisman, Rev. Sci. Inst. **3**, 7 (1932).

[11] Zisman and Yamins, Physics **4**, 7 (1933).

[12] Yamins and Zisman, J. Chem. Phys. **1**, 656 (1933).

trode to vibrate 250 to 500 times per second instead of coating it with polonium. This method of the vibrating condenser seems to give the same results as that of Guyot and Frumkin, but thus far it has been tested only with respect to the general order of magnitude of the values for the condensed film.

II. Apparatus and Preparation of Films

The film balance, of the type designed by Adam[13] and modified by Freud and Harkins[14]

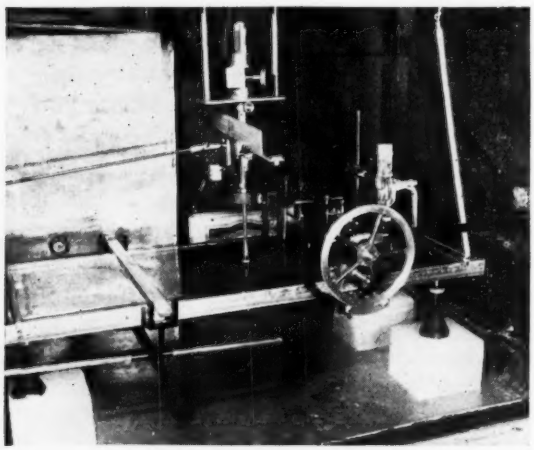

Fig. 1. Apparatus for the simultaneous measurement of film pressure and surface potential, and for the determination of the location and contour of continents and islands of condensed film. This view is taken inside the large thermally insulated metal box which surrounds the apparatus. The Compton electrometer is inside the smaller lead box in the left-rear.

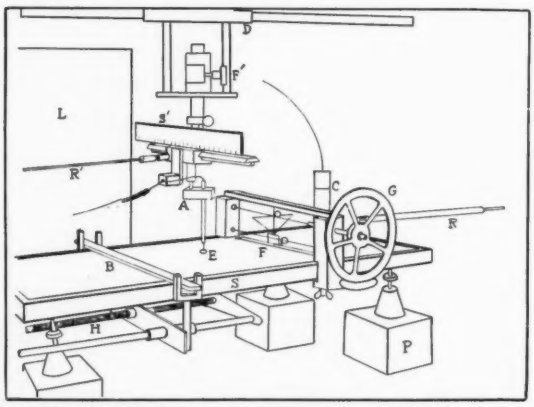

Fig. 2. Diagram of apparatus shown in Fig. 1.

[13] Adam, *Physics and Chemistry of Surfaces*, Clarendon Press, 1930.

[14] Freud, Doctoral Dissertation, University of Chicago, 1927.

(Figs. 1 and 2), is supported on four blocks of paraffin, P, inside a large grounded metal box supported on a heavy concrete pier. Heat insulation of the box is provided by a thick layer of wool felt on the outside. Inside this box is a smaller box L, the inner walls of which are heavy sheets of lead. This box holds the sensitive Compton electrometer used to determine the potential. The electrical circuit is shown in Fig. 3. This is so

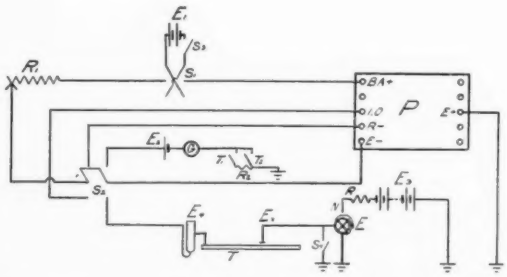

Fig. 3. Diagram of electrical connections for determination of film potential. T, trough of paraffin or brass; P, potentiometer; E_4, one to six calomel cells; E, Compton electrometer; E_1, battery for potentiometer; E_2, standard cell, E_3, battery for charging electrometer; S_2, double-throw double-pole switch: upper position for setting R_1 with P at zero; lower position for measuring E_x on potentiometer with zero reading for E.

arranged that the potential may be either given by a potentiometer, or read directly from the electrometer deflection, which amounts to 680 mm per volt at a scale distance of about 1.5 meters. A much greater sensitivity is easily obtained.

The front of the large box is occupied by two glass doors with metal frames, but the inside of the glass is covered with earthed copper gauze.

As has been customary in this laboratory for many years, the entire operation of the film balance system, except for filling the trough and spreading the film, is carried out by means of handles outside the box. This procedure was also used with the electrical system. Thus the handle on the ebonite rod R controls the film balance, while that connected with the screw H controls the barrier B by means of two forked supports, one at each end of B. A handle on the rod R' serves to shift the position of the silver electrode, coated with polonium, from side to side with respect to the surface of the solution, while the carriage D is moved lengthwise of the film by another rod and handle not shown. Lengths along the

trough are measured by the scale S and those across it by S'. The electrode E is insulated by the amber supports A and its height is adjusted by the rack and pinion F'.

The float F of thin sheet metal is attached to the end supports of the balance by vertical sheets of thin gold shaped as the letter S. The graduated circle G is read by a vernier V below it. The trough T is 25 cm wide and 60 cm long. The surface back of the float F is kept clean by sweeping with various barriers which are not shown. Troughs of glass, paraffin, and of brass carefully coated with paraffin, were found to give results which agree remarkably well.

Preparation of films

The organic acids used for the preparation of the films were the purest that could be obtained. Each sample was washed with a large quantity of pure hot water, and was then recrystallized twice from pure alcohol, and finally dried carefully.

A dilute solution of the film forming substance in a low boiling petroleum ether was prepared. A portion of this solution was put in a special weight pipette designed by Harkins and Freud in 1926. The carefully ground caps and stopper of these pipettes must be kept dry, and if this is done they prevent the evaporation of the solution during the process of weighing. Thus the error of 10 percent in the work of Schulman and Rideal caused by the determination of the amount of solution by counting the drops was avoided. Adam and Harding and Schulman and Hughes have greatly reduced the error below that of Schulman and Rideal by the use of the volumetric pipettes suggested by Fosbinder.

It is obvious that with dilute films the distribution of the islands is highly dependent upon the way in which the drops of the dilute solution are placed upon the surface of the liquid in the trough. An extensive study was made of the effects on the surface potentials of dilute and condensed films as affected by various methods of putting the dilute solution upon the surface.

Electrodes

Many experiments were carried out to determine the best electrode to use with a radioactive substance. A platinum gauze electrode in the form of a small basket with from six to twenty radon tubes laid directly on the metal was found to give potentials which varied (1) with the number of radon tubes, (2) with the distance of the electrode above the surface of the solution, and (3) with the previous treatment of the platinum. For example, the potential in millivolts against the surface of 0.1 molal hydrochloric acid was 95 after immersion of the platinum in cold cleaning solution, -200 after cathodic electrolysis, 312 after anodic electrolysis, and 520 after heating to redness in a Bunsen flame. After standing for several hours the potentials of the electrodes listed above became more nearly the same, but constant results could not be obtained.

Electrodes of various metals covered with polonium were much more satisfactory. Wires and disks of platinum, cadmium, zinc, tin, gold, copper and silver were tried. Good values were obtained with copper but even more constant values with silver, so this metal was used in all of the experiments reported.

Connection with the solution was made by from one to six calomel half cells, which seemed to give better results than a large silver-silver chloride electrode.

III. Potential and Force-Area Values for Films of Myristic Acid

Interesting results were obtained with monomolecular films of myristic acid on 0.01 N hydrochloric acid (Fig. 4), since the curves which give the variation of potential with the mean molecular area of the myristic acid are much more simple than those of Schulman and Rideal.

At molecular areas from 15 to 52 sq. Å the values of the surface potential are represented by a moderately smooth curve. If any inflection at all appears at 32 sq. Å, which is the location of a marked inflection in the pressure-area (FA) curve, it is a very minor one. At areas above 52 sq. Å the pressure falls to that of the gaseous film. The four sets of measurements at areas higher than this are of considerable value in connection with an interpretation of the characteristics of the film. In experiment 1, represented by circles with the lower half inked in, the film was at the beginning very dilute. As the film was compressed from high areas down to 57 sq. Å the potential

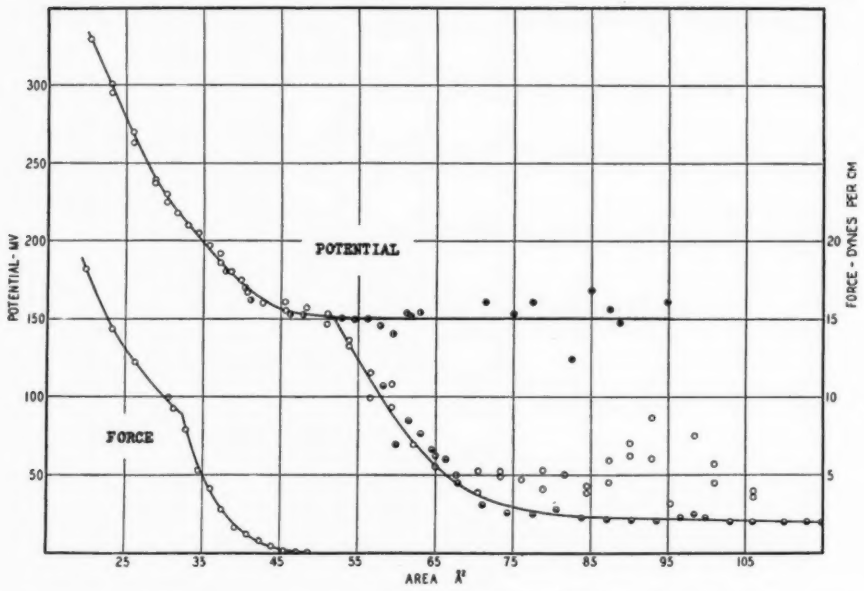

FIG. 4. Lower left = pressure-area relations. Upper right = potential-molecular area relations. Films of myristic acid at 17°. The surface potentials are variable at molecular areas greater than 52 Å. The lower curve corresponds to the potential of a gaseous film, and the higher curve to that for an island.

values increased along a moderately smooth curve. At 115 sq. Å the surface potential was 19 mv, and at 65 sq. Å about 60 mv.

In experiment 2 (open circles), the potentials for areas below 65 sq. Å agree moderately well with those for experiment 1, but at higher areas they vary between 30 and 90 mv. In experiment 3 (black circles) the potentials are much higher, and between 150 and 170 mv. In experiment 4 (left half of circle open) the range of values was more limited, but of the same order as for experiment 4.

In experiments 3 and 4 the silver-polonium electrode was directly over a continent or island of film in what has been called the "expanded" state, which represents a moderately closely packed film. In experiment 1 at 115 sq. Å molecular area the electrode was largely over an area of gaseous film, but the potential was raised by adjacent continents or islands. As the area was decreased the area of gaseous film under the electrode was at first decreased only very slowly, but after the molecular area had decreased to less than 75 sq. Å the continents of concentrated film occupied such a great fraction of the area that further decrease of area caused the gaseous film region to disappear rapidly, with a resultant rapid increase of potential. However, when the whole film became of the "expanded" rather than of the gaseous type, the rate of increase became much more slow, since the gaseous phase was no longer present, and thus could not disappear.

At areas below 52 sq. Å the pressure was above that of the gaseous film, and the continents of concentrated film were pressed together, so below this area a single curve is sufficient to represent the potential values. However, all irregularities do not always disappear, since the effect of the contour of the continents is not entirely effaced. Between 53 and 22.5 sq. Å for the molecular area the results given in the figure are those of experiments 2 and 3, and are represented by a single smooth curve. Curves of this type were obtained in a considerable number of similar experiments between 10° and 30°. The potential exhibited a general decrease as the temperature was increased. In a few experiments a minor inflection occurred at the same area as that at which the marked inflection occurs in the force-area curve.

A geographic plot of the variation of the surface potential for films of myristic acid at 23° is given in Fig. 5. The variation in potential at the different areas was as follows:

Molecular area in sq. Å	56.5	37.1	25
Minimum potential in mv	0.0	155	238
Maximum potential in mv	176	175	281

In general the variation of the potential decreased as the area decreased toward 20 sq. Å. At 12° Adam and Harding find a variation of 166 mv at a molecular area of 70.6 sq. Å, with a minimum of 1 mv and a maximum of 167 mv.

Palmitic Acid

Area

48 Å²

Barrier→

68	217	315	330
35	59	141	200
233	130	50	61

←Float

282.2 Å²

345	315	324	315
297	310	311	312
319	314	308	305
331	313	313	331

Pentadecylic Acid

48.5 Å²

155	163	170	175
160	20	165	170
165	24	101	170

39.7 Å²

179	183	183	180
179	180	181	180
175	179	180	179
175	179	181	186

Myristic Acid

56.5 Å²

10	22	31	42	90	165	158
0	15	30	38	164	160	156
5	35	51	165	167	165	165
48	98	166	174	171	176	164
145	148	170	169	170	103	199

37.1 Å²

175	170	170	160	157
172	165	168	160	155
169	166	167	164	158
166	169	188	165	155
163	167	190	195	—

25.0 Å²

255	248	255	281
255	248	255	245
255	252	258	243
254	277	261	238
253	254	261	256

FIG. 5. Distribution of surface potential values according to the location of the electrode over the film. The higher values correspond to the more concentrated portions of the film. Barrier on the left, and float attached to film balance on the right. Palmitic acid film at 23°C. Other films at 17°C.

IV. Islands and Other Non-Homogeneous Areas in Films

One of the first steps in the work reported in the present paper was to study the effect on the surface potentials at different positions of various methods of dropping the dilute solution of the film forming substance on the surface. For this purpose the silver-polonium electrode was mounted in such a way as to be easily moved in either direction from barrier to float. It was found that when the potential was plotted against the distance a high peak or a wide plateau appeared near the position where a drop had been placed, provided the measurements were made soon after the drops had been deposited. Often, however the islands move quickly from their original positions. The best results are obtained with large islands anchored by some solid object.

The preceding section illustrates the existence of such islands in gaseous films of myristic acid. These appear as floating independent islands at molecular areas higher than those given in Fig. 5, where the islands have largely merged into continents. At extremely great molecular areas the islands disappear or become less prominent.

With a film of palmitic acid at 23° and a molecular area of 48 sq. Å, the surface potential varied from 35 to 330 mv, while at 28.2 sq. Å the potential had become more constant and between 297 and 345 mv. That the potential is not more variable than this is surprising, as the film pressure is still very low. At lower areas and higher pressures the potential becomes still more constant.

With pentadecylic acid the film pressure rises at much higher molecular areas, so that, while at 48.5 sq. Å the potential varies from 20 to 175 mv, at the somewhat high area of 39.7 sq. Å most of the variation has disappeared and the potential varies only from 175 to 180 mv, except in one corner of the film where one value was obtained at 185 mv.

At lower areas than those given in Fig. 5 the potentials in general become more constant and except for some peculiar positions, such as a corner of the film, do not vary more than 3 to 5 mv. However, at molecular areas much below 20 sq. Å the potential again becomes more variable, and in general decreases, since most films crumple at these lower areas, and regions

are produced in which the molecules of the film do not preserve their regular orientation with respect to the surface of the liquid.

V. General Potential Relations for Films of Fatty Acids

The surface potential of a fatty acid film increases rapidly with the length of the hydrocarbon chain. Thus at 20 sq. Å the potentials for stearic (18), palmitic (16), pentadecylic (15), myristic (14), and lauric (12) acids, are respectively 400, 383, 375, 336 and 276 mv.

Each curve in Fig. 6 represents a single experiment. At areas greater than that at which, according to the force-area curves, the pressure apparently falls to zero, other experiments gave other results, but at smaller areas there was good agreement.

The presence of the double bond of oleic acid produces a lowering of about 200 mv, a surprisingly great lowering of the surface potential. The presence of the single double bond lowers the potential to about the same extent as the removal of six carbon atoms.

The values for myristic acid (14 C) given in Fig. 4 lie below those for pentadecylic acid (Fig. 6). The flat portion of the curve lies at about 150 mv for the former, and 170 mv for the latter.

VI. Alcohol Films

The two higher alcohols investigated (Fig. 7) show the same relation as the acids in that the potential is higher for the alcohol with the longer chain. The surface potential is, however, higher for the alcohols than for the corresponding acids.

VII. Paraffin Layers

Nonpolar hydrocarbons on a clean water surface exhibit different potential relations from a fatty acid. Thus when paraffin was spread from a solution in ligroin, the potential of the interface became 60 mv more positive with respect to the clean surface when the area was decreased to one-fourth the initial area. At the same time the pres-

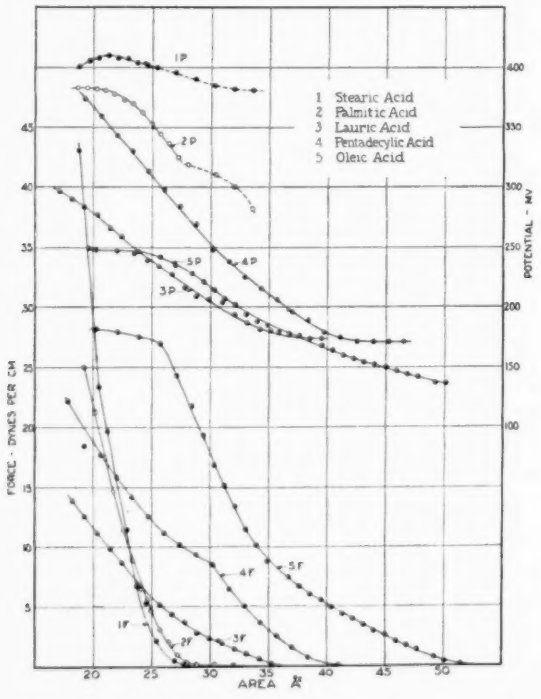

Fig. 6. Surface potentials (P) and film pressure values (F) for films of organic acids on 0.01 molal hydrochloric acid. The points plotted correspond to individual experiments. The potentials for areas greater than that at which the special force-area curve apparently merges with the x-axis, are uncertain. Temperature, 17°.

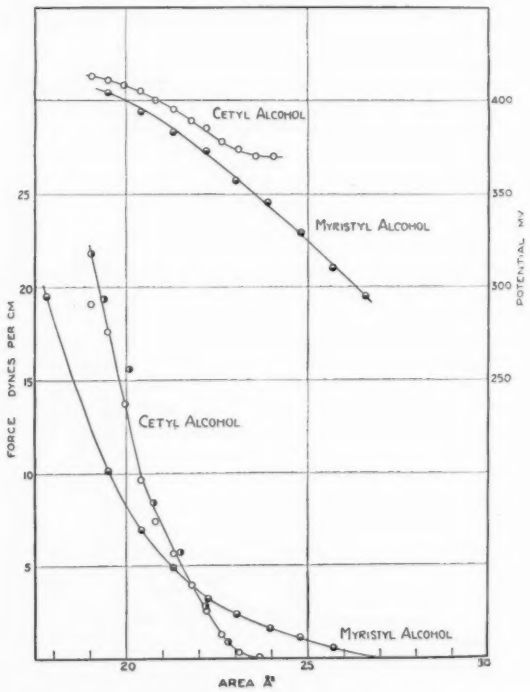

Fig. 7. Surface potentials (above) and surface pressures (below) for films of normal alcohols. As with acids the potential increases with the length of the carbon chain. Temperature, 17°.

sure exerted on the float was of the order of one dyne per cm. The film of paraffin was visible to the eye.

VIII. Surface Potentials, Dipole Moments, and the Orientation of Molecules in Surfaces

The surface potential of a film, defined as the contact potential of the surface covered by the film minus that of the clean surface of the underlying phase, is undoubtedly related to the dipole moments of the film producing substance, to the extent of the orientation of the molecules, and to other factors, some of which are geometric.

The orientation of the molecules in a surface was supposed by Hardy[15] to be due to the unsymmetrical field of the molecules, and Harkins, Davies and Clark[16] took a somewhat similar view when they assumed that the potential differences at surfaces, as well as the other energy relations may be explained on the basis of the polar setting of molecules in surfaces. An addition to this electrical theory of surfaces was made by Harkins and King[17] who consider especially the relations between surface energy, cohesion, and the polarity of the molecules which constitute the legends or liquids involved.

In a monomolecular film on water, on any other liquid, or on a solid, the molecules in the film, and to some extent the adjacent molecules of the surface region, orient themselves in such a way as to decrease the free surface energy.

In this discussion it will be assumed that just as a stone in a position above the surface of the earth possesses an energy of position, usually designated as its potential energy, a molecule in the surface possesses an energy of position which will also be termed its potential energy. The potential energy of the weight may be referred to an arbitrary zero position at the level of the earth's surface, while that of the molecule is referred to a position in the liquid phase underneath the surface region. The work done in lifting the weight is done against the force of gravitation, while that of lifting the molecule is opposed by the cohesional forces which may also be considered as the resultant of the electrical forces.

The potential energy of the stone with reference to its zero position, may be determined by lifting the stone, initially at rest, to the specified position. The potential energy is said to be equal to the work done. If the stone, when at the level of the earth's surface, had a velocity in the upward direction, then the work done would be less, but the final energy of position at a specified height would not be changed.

A molecule inside the liquid can in general move into the surface region only when its energy of position, which we will here call its potential energy, is supplied in the form of work, heat, or of some other class of energy.

The total energy necessary to form one sq. cm of surface is supplied partly as work γ and partly from the kinetic energy of the molecules themselves. The molecular kinetic energy is the source of almost the whole latent heat l of the surface. Thus a surface has a lower temperature than the liquid when it is first formed.

The value of the heat function h for unit area of a surface is given by the relation

$$h=\gamma+l. \tag{1}$$

In an isothermal change h is the total amount of energy which must be added to the system when a unit area of surface is created, and is usually designated as the total surface energy.

It will be assumed, on the basis of our general knowledge of surface energy relations that π, the sum of the potential energies of all of the molecules in 1 sq. cm of the surface region is very nearly equal to h the energy supplied to form unit area of the surface, or

$$\pi=h. \tag{2}$$

The orientation of the molecules causes a decrease in the free energy (γ) of the surface, while on the whole the latent heat is not greatly changed, so the total surface energy h is decreased by the orientation. Thus the sum (π) of the potential energies of the surface molecules is also decreased (Eq. (2)) by the orientation below that for a random distribution. Since the potential energy of the molecules in the surface is assumed to be due to electrical forces, it may be called the total electrical energy.

[15] Hardy, Proc. Roy. Soc. **A86**, 634 (1911–12).

[16] Harkins, Davies and Clark, J. Am. Chem. Soc. **39**, 541 (1917).

[17] Harkins and King, J. Am. Chem. Soc. **41**, 970 (1919).

The mutual energy due to the interaction of the dipole moments of the molecules constitutes only a part of this electrical energy. In the older theory the remainder of this energy is attributed to quadrapoles or multipoles but more recently the theory has been modified by the introduction of some of the principles of quantum mechanics, as, for example in papers by Lennard-Jones.[18]

While thermodynamics predicts only the decrease of free surface energy as associated with the molecular orientation, the above discussion seems to indicate that this is in general accompanied by a decrease in the total electrical surface energy. While it is not from these relations certain that the electrical energy of the dipoles must decrease with the decrease of total electrical energy, the dipoles make an important contribution to this energy, so it seems probable that such a decrease occurs. Also from the known behavior of large scale dipoles which are free to change their orientation and position it may be deduced that the molecular dipoles will shift and orient themselves in such a way as to decrease their mutual electrostatic energy. However, the decrease may not be to a minimum since it is possible that there may be in some cases a resultant increase of the energy of the multipoles. The dipoles involved include those of the liquid (usually water) as well as those of the organic molecules of the film.

The potential difference (ΔV) between two parallel plane metal plates of an ordinary parallel plate condenser, *with a unit dielectric constant, is*

$$\Delta V = 4\pi\sigma s, \tag{3}$$

in which σ is the surface density of the charge and s is the distance between the plates.

Since $\sigma = ne$, where n is the number of electronic charges per sq. cm and e is the electronic charge,

$$V = 4\pi n(es) = 4\pi n\overline{m}, \tag{4}$$

in which $\overline{m}$ is the product of the value of the electronic charge by the distance between the plates.

Schulman and Rideal[19] consider that this equation, which they attribute to Helmholtz, may be applied to calculate the dipole moment of the molecules of the film on the basis that $\mu_s = \overline{m}$ if n is considered as the number of molecules in one sq. cm of film. Thus

$$\mu_S = \Delta V/4\pi n. \tag{5}$$

They find, however, that the value thus obtained is only about one-sixth, or even less, of that determined by the standard methods. Their explanation of the discrepancy may be found in their papers.

Our point of view is different in that we consider that the equation of the parallel plate condenser is not valid for the case of a film.

(1) The equation assumes that the charges are on metal plates, which is far from true, since they are assumed to be associated with the molecules.

The correct equation for parallel plates of an insulating material is

$$\Delta V = 2\pi n\overline{m}'. \tag{6}$$

If now it is assumed that $\mu_s = \overline{m}'$, then the value of μ_s thus obtained is, obviously, much closer to the standard value of the dipole moment. However, this does not show that Eq. (6) is suited to the calculation of the dipole moment.

(2) The Eqs. (5) or (6) were developed for a large scale condenser. There is no evidence that they are valid for a condenser in which the charges are attached to molecules, and are a distance apart which is of the atomic scale.

(4) The charges of this atomic scale condenser are only imaginary or apparent charges, though the dipole moment is real.

(5) Even if these charges are assumed to be real and to have a definite position, the charges would move with the heat vibrations of the molecules in such a way that any instant both positive and negative charges would probably be on both sides of a plane which represents the height of the mean position of all of the imaginary charges, both positive and negative. Even if this is not the case the charges do not lie in a plane.

(6) Both the molecules and the positions of the imaginary charges are in rapid motion, which is not taken into account in the development of the equation.

Thus μ_s does not seem to be a dipole moment, but is a quantity which may be found useful,

[18] Lennard-Jones, Proc. Phys. Soc. (London) **43**, 461 (1931).

[19] Schulman and Rideal, Proc. Roy. Soc. **A130**, 260 (1931).

since it is equal to the surface potential divided by the number of molecules in unit area of the film and this quotient multiplied by the constant $1/(4\pi)$. It seems preferable to make the constant unity. This gives

$$m = \Delta V/n. \tag{7}$$

Either m or μ_s is thus related in a simple way to the surface potential and the concentration of the molecules in the film, and in a complicated and unknown way to the standard dipole moment. The values of m are given in Table I.

TABLE I. *Values of $m = V/n$ for films of fatty acids and alcohols, as calculated from surface potentials.*

Substrate = 0.01 N HCl (for the lauric acid films this was nearly saturated with NaCl to reduce the solubility of the lauric acid).

Compound	$m \times 10^{18}$ volts molecule^{-1} cm^{-2} at area 20	25	30	40	48	Sq. Å
Lauric acid	1.9	2.0	2.0	2.2_6		
Myristic acid	2.2_6	2.3	2.3	2.3	2.4	
Pentadecylic acid	2.4	2.6	2.5_4	2.3		
Palmitic acid	2.5_4	2.9_4				
Stearic acid	2.7	3.3				
Oleic acid	1.6_5	2.0	2.1_6	2.2	2.2_5	
Myristyl alcohol	2.6_6	2.7				
Cetyl alcohol	2.9					

In the determination of the factors which are of the most importance in their effect on the surface potential, the variation of this potential with temperature is of prime importance. In the experimental work the surface potential was found to usually exhibit a decrease with rising temperature, but a rise in temperature often changed the state of the film. The effect of increase of temperature on the surface potential of condensed films will be studied more carefully later.

IX. General Discussion

There are certain features of the potential-area curves which are of importance. There is no indication in these curves of the numerous "surface phases," such as the "homalic" and "prehomalic" phases of Schulman and Rideal. Since these "phases" are not indicated by any known force-area determinations it is evident that there is no evidence that such phases have any existence, especially since both Adam and Schulman and Hughes have also failed to find them. Doubt is also thrown upon the value of the consideration of condensed and expanded films, as defined by Adam, as distinct phases. If films in general, after being compressed to a molecular area of about 20 sq. Å, are subjected to a constantly reducing pressure, it is found that there is considerable hysteresis, both with respect to pressure and potential. With increasing pressure there is in general no special variation in the curvature of the potential-area curve at the point where the force-area curve indicates what has been considered to be the transition point between expanded and condensed phases. This fact seems to indicate that there is no sharp change in orientation or structure in the film at this point. The above considerations do not show that the terms "condensed" and "expanded" with reference to films should be abandoned, but they indicate that there is a need for a more detailed study of the reversibility of the supposed transition between the two states.

As has been stated earlier the curves for the potential-area variations as given in this paper, represent individual sets of determinations. These have been checked by numerous sets which have not been reported. The curves presented are not the smoothest which were obtained, but they were taken from several sets in good agreement with each other. The curves exhibited by Adam and Harding, and by Schulman and Hughes, in recent papers are in general either averaged or idealized, so that no specific data are presented. In some cases the same substances were used by Adam and Harding and in the present investigation, with no apparent disagreement in the results. However, since our temperatures were higher the potentials obtained are lower.

The specific experiments on films of organic acids, as presented in Fig. 6, exhibit several features of general interest. For example the surface potential for films of stearic acid at areas greater than 27.5 sq. Å may be as high as those given by the curve, or lower values may be found. However, in this individual set of determinations the potential-area curve gives no indication of any transition whatever. The same is true of pentadecylic films above 40 sq. Å. Of the two condensed films that of stearic acid is much more

coherent than that of pentadecylic acid. With such highly coherent films as that of stearic acid the surface potential is lowered only slightly as the pressure on the film is reduced. Thus it appears that the molecular orientation in such a portion of a solid film is not materially affected by the release of pressure on the film to almost a zero value. If time is given for a two dimensional evaporation of the film substance, at extremely low pressure, the surface potential is reduced to a few millivolts.

The surface potential is found to vary with

1. the length and shape of the hydrocarbon part of the organic molecule,
2. the nature and number and position of the polar groups in the molecule,
3. the hydrogen ion concentration of the solution,
4. the temperature,
5. the concentration of the film forming substance in the film,
6. the nature of any ions of salts present in the solution,
7. time, including phenomena of hysteresis.

With changes of acid concentration between 0.01 and 0.1 N the surface potentials for films of acids and alcohols was found to be only slightly affected.

The effect of hydrogen ion concentration may be illustrated by experiments on polymers of hydroxydecanoic acid with molecular weights from 208 to 16,900 and general formula $HO[(CH_2)_9COO]_xH$. A change from 0.01 normal hydrochloric acid to the same concentration of sodium hydroxide increased the molecular area in the condensed film by about 30 percent and decreased the surface potential by ten to twenty percent. Proteins were even more largely affected by the hydrogen ion concentration, as corresponds with the work of Gorter. The details of the effects of hydrogen ion concentration upon the areas and surface potentials of organic substances of high molecular weight will be left to another paper.

In order to test the assumption presented in Section I of this paper, that the method used in this work is suitable for the determination of either the surface potentials of solutions or the contact potentials of metals, the carefully cleaned surfaces of several metals were substituted for the liquid surface. The differences in these values were found to give the contact potentials of the metals as accurately as they are known for like conditions. It is obvious that the calomel cell is eliminated from the circuit for such a measurement.

The writers wish to thank Dr. Julius Stieglitz, and the Chemical Foundation, whose aid has made this work possible.

Acoustical Studies. II.[1] The Behavior of a Gas with Several Independent Internal Energy States

WILLIAM T. RICHARDS,[2] *Princeton University*

(Received August 24, 1933)

The method founded by Einstein for the description of the acoustical behavior of a dissociating gas has been extended to cover a nondissociating gas in which five groups of internal energy states have different relaxation times. Since the resulting expressions are unwieldly, approximations based on them, which permit the rough description of experiment, have been given. The study of the variation of the velocity of sound with frequency demonstrates differences in the relaxation times of the various states. The variation of the velocity of sound with pressure shows in addition effects due to three-body collisions or to radiation from optically active states. It has been suggested to treat the temperature coefficient of the kinetic relaxation time by the introduction of an empirical quantity which may be called "the activation energy of collision." By this means a rough kinetic analysis of the transition probability is possible. In gaseous mixtures, the relative spatial and energetic effectiveness of various types of collisions in exciting internal energy may be compared.

INTRODUCTION

GREAT progress has been made during the past few years in the measurement of the velocity of sound in gases. Because of the invention of the sonic interferometer by Pierce[3] it is now possible to determine this quantity in a gas of any desired purity with an accuracy of 0.1 percent from audible frequencies to frequencies somewhat above a million cycles per second. There is every reason to suppose that, with sufficient incentive, even this accuracy may be considerably improved. Similarly, interferometric absorption measurements, particularly owing to the investigations of Hubbard,[4] are rapidly becoming capable of exact interpretation. It is worth while, therefore, to discuss in detail the information which may be expected to become available from extremely accurate acoustical measurements.

Two methods have been used to describe the propagation of sound in acoustically dispersive gases. The first of these evaluates the absorption and dispersion by calculating as a function of frequency the number of molecules making the transition between two given energy states under the influence of the adiabatic acoustical cycle. It was founded by Einstein,[5] who considered the case of a dissociating gas, and has recently been extended by Kneser[6] and by Rutgers[7] to include nondissociating gases. The second method uses the conception of the fictional temperature of an energy state or group of energy states, and calculates as a function of frequency the relation between this fictional temperature and the temperature of the translational terms in the acoustical cycle. It was first published by Herzfeld and Rice[8] and was generalized immediately thereafter by Bourgin[9] to cover gaseous mixtures and effects due to radiation. The two methods are equivalent descriptions of the same process, and with similar postulates must yield similar conclusions. They have converged upon

[1] It is considered that the communications of Richards and Reid, J. Chem. Phys. **1**, 114,737 (1933), are Parts I and Ia of this series.

[2] Fellow of the John Simon Guggenheim Memorial Foundation.

[3] Pierce, Proc. Am. Acad. Sci. **63**, 1 (1928).

[4] Hubbard, Phys. Rev. **41**, 833 (1932), and earlier papers there cited.

[5] Einstein, Sitz. Ber. Akad. 380 (1920).

[6] Kneser, Ann. d. Physik **11**, 761 (1931); **16**, 337 (1933).

[7] Rutgers, Ann. d. Physik **16**, 350 (1933).

[8] Herzfeld and Rice, Phys. Rev. **31**, 691 (1928). Among the numerous and important conclusions reached in this paper is included a demonstration that absorption due to viscosity and heat conduction plays a minor rôle if a heat capacity effect is also present. This is formal justification for disregarding this quantity in using Einstein's method.

[9] Bourgin, Nature **122**, 133 (1928); Phil. Mag. **7**, 821 (1929); Phys. Rev. **34**, 521 (1929) and **42**, 721 (1932).

expressions for the velocity and absorption of sound which are practically identical, and are suitable for the description of experiment if the dispersion is produced by the gradual disappearance from the heat capacity at high frequencies of the transition from the zero point energy state of a vibration to its first excited state.

The purpose of the following discussion is to amplify these expressions to include as many transitions as may practically be necessary. The criterion of practical necessity is the calculation of the velocity of sound at very low frequency within limit of experimental error from the same relationship which is used to describe the dispersion. The method of Einstein has been used in preference to that of Herzfeld-Rice-Bourgin because of its greater simplicity. It must again be emphasized that either method is suitable for the purpose, and that the description of cases of any desired complexity is merely a matter of industry. Sections 1, 2, 3 and 4 of the following study are devoted to the development of expressions for the velocity and absorption of sound suitable for use with gases of increasing numbers of frequency dependent internal energy terms. Section 5 considers the bearing of the earlier sections on experiment, and shows how suitable measurements may lead to an increased knowledge concerning collisions of the first and second kinds. An effort has been made to make Section 5 intelligible to a reader who does not wish to follow the mathematical detail of the earlier Sections. Finally, in Section 6, gaseous mixtures and more complex transitions in pure gases are briefly considered.

1. The Method of Einstein

If we consider the propagation of sound in a pure nondissociating gas which has a molecular heat capacity at constant volume greater than $(3/2)R$, we may expect with increasing frequency one of two alternative types of behavior. Either the rotational and vibrational energy will be brought into equilibrium by each collision, in which case the gas will not become acoustically opaque until frequencies of the order of magnitude of the collision frequency have been reached, or the adjustment of rotational and vibrational energy may be hindered by some spatial or energetic property of the colliding molecules, in which case it will lag behind the translational adjustments and the gas will become partially opaque at lower frequencies. It is with the second of these possibilities which we are here concerned. If the internal heat capacity is due to changes between many states, each of these changes may be characterized by a separate rate of transition. The method of Einstein may be adapted to supply a quantitative description of the rates of these transition processes.

If π is the excess pressure in an element of volume of the gas at a distance x from the source, and Δ is similarly the density increment of this element, one may write for the time variations of these quantities

$$\pi = \pi_0 e^{i\varphi} e^{i\omega(t-ax)} \tag{1.1}$$

and

$$\Delta = \Delta_0 e^{i\omega(t-ax)}, \tag{1.2}$$

where ω is the frequency of sound in radians, and t refers to time. The velocity and absorption of sound are related to a by the expression

$$a = (1/V - i\beta/\omega) \tag{1.3}$$

if V is the phase velocity ($2\pi V = \omega\lambda$) and β the absorption coefficient for the amplitude expressed as a reciprocal distance. The propagation of sound in an absorbing medium is therefore in close analogy to the propagation of light in a medium with a complex refractive index. The angle φ expresses the phase lag between pressure and density increments[10] and characterizes both absorption and dispersion. From the real and imaginary parts of (1.1) and (1.2) and the equation of motion of a plane compressional wave it is found that

$$\cos\varphi = (\pi_0/\Delta_0)(1/V^2 - \beta^2/\omega^2) \tag{1.4}$$

and

$$\sin\varphi = (\pi_0/\Delta_0)(2\beta/\omega V). \tag{1.5}$$

Einstein has demonstrated that the quantity π/Δ corresponds to the square of a complex velocity of sound, and from this it follows that

$$\text{r.p. } \pi/\Delta = (\partial P/\partial\rho)_{\text{adiab.}} = V^2\mu \tag{1.6}$$

[10] Kneser, second reference under footnote 6, has given an illuminating specific picture of this lag and the irreversible energy change it implies.

and that

$$\text{i.p. } \Delta/\pi = -i2\beta/\omega V = -ih/2\pi V^2, \quad (1.7)$$

when h, in (1.7), designates the absorption coefficient for the intensity expressed as a reciprocal number of wave-lengths. The factor μ in (1.6) is

$$\mu = (1-h^2/16\pi^2)/(1+h^2/16\pi^2)^2 \quad (1.8)$$

and may be taken as unity except in the neighborhood of the absorption maximum.

In general there will be absorption only when the dispersion $dV/d\omega$ is appreciable.[8] It is difficult, however, to formulate the relation between the two, since it depends upon the characteristics of π/Δ. It is more convenient, therefore, to compare absorption and velocity, which are more simply related as in (1.4). This may be expressed by combining (1.4) with (1.6) to give

$$(\pi_0/\Delta_0)(1+h^2/16\pi^2) = V^2. \quad (1.9)$$

The complex velocity of sound and the properties of the gas are related by the *characteristic acoustical equation*. This may be written

$$\pi/\Delta = (P/\rho)(1-\Delta Pv/P\Delta v) \quad (1.10)$$

if P and ρ refer to pressure and density respectively, and ΔPv and $P\Delta v$ are complex energy increments which vary with time about their respective amplitudes by a factor $e^{i\omega t}$, and which are out of phase in the dispersive region of frequency. The problem of calculating the acoustical behavior of a gas is therefore the problem of evaluating ΔPv and $P\Delta v$ in terms of pressure, frequency, temperature, and quantities descriptive of rate of energy adjustment and heat capacity. This is conveniently effected by means of the equation of state and the equation for adiabatic change in the gas, the equation for adiabatic change being related to frequency by means of a series of reaction equations.

In finding ΔPv the gas will be considered ideal, and effects due to departures from ideal behavior later discussed. For an ideal gas the *displaced equation of state* is

$$\Delta Pv = R\Delta T, \quad (1.11)$$

where R and v are considered gram molecular quantities, and ΔT refers to a temperature displacement of the translational terms. A relation between ΔT and the displacement of the internal energy is given by the *equation for adiabatic change*, which may be written

$$-P\Delta v = (3/2)R\Delta T + E_1 \sum_k k\Delta n_k + E_2 \sum_l l\Delta n_l + \cdots. \quad (1.12)$$

The first term on the right refers to changes in translational energy, and the subsequent terms express energy changes in sets of rotational and vibrational states corresponding to the frequencies $\nu_1 = E_1/hN_A$, $\nu_2 = E_2/hN_A$, etc., h being Planck's constant and N_A Avogadro's number. The symbols n_k, n_l, etc., refer therefore to fractions of gram molecular weights in the kth and lth, etc., states, and are subject to the conditions $\sum_k n_k = 1$, etc., and $\sum_k \Delta n_k = 0$, etc., owing to the law of the conservation of matter. The actual number of molecules per cubic centimeter in the kth state is Nn_k, if N is the total number of molecules per cubic centimeter. If C_1 is the equilibrium heat capacity due to the first set of states, and ΔT_1 its fictional temperature displacement in the sound wave, $E_1 \sum_k k\Delta n_k = C_1 \Delta T_1$ where $\Delta T \geqq \Delta T_1 \geqq 0$. The calculation of the velocity and absorption of sound in a pure gas of given properties follows from an appropriate development of (1.12) on the basis of a reaction equation, and the subsequent substitution of (1.11) and (1.12) in the characteristic acoustical equation (1.10).

In Sections 2, 3, and 4 of this communication the characteristic acoustical equation of the gas will have the form

$$\frac{\pi}{\Delta} = \frac{P}{\rho}\left\{1 + \frac{R}{C_\infty + Rx^2 n_0[(a+ib)/(c+id)]}\right\}, \quad (1.13)$$

where C_∞ is that part of the heat capacity which is independent of frequency in the range considered,

$x=E/RT=h\nu/kT$, and the constants n_0, a, b, c, and d vary with the complexity of the internal energy of the gas. According to (1.6), therefore, the velocity is given by

$$V^2\mu=\frac{P}{\rho}\left\{1+R\frac{(C_\infty+Rx^2n_0a/c)+(C_\infty d^2/c^2+Rx^2n_0bd/c^2)}{(C_\infty+Rx^2n_0a/c)^2+(C_\infty d/c+Rx^2n_0b/c)^2}\right\}=\frac{P}{\rho}\gamma_\omega, \tag{1.14}$$

where γ_ω may be called the "apparent heat capacity ratio" at the frequency ω, and the dimensionless absorption coefficient, from (1.7) to the first order of small quantities is

$$h=2\pi\left\{\frac{(C_\infty+Rx^2n_0a/c)(C_\infty d/c+Rd/c+Rx^2n_0b/c)-(C_\infty d/c+Rx^2n_0b/c)(C_\infty+R+Rx^2n_0a/c)}{(C_\infty+Rx^2n_0a/c)(C_\infty+R+Rx^2n_0a/c)+(C_\infty d/c+Rx^2n_0b/c)(C_\infty d/c+Rd/c+Rx^2n_0b/c)}\right\}. \tag{1.15}$$

Suitable constants for substitution in (1.14) and (1.15) will now be evaluated.

2. A Gas with Three Sets of Energy States

The simplest case of practical interest is that in which the dispersive region results from a lag in the adjustment of a single quantum transition, as, for example, the transition from the zero point state to the first excited state of a deformation vibration. To describe this various parts of the energy may be considered as contained in three types of energy states. The states E_{I} are the translational states. The states E_{II} contain in addition all rotational and vibrational states in their equilibrium distribution with the exception of a single mode of vibration, which is considered to have its zero point energy. E_{III} differs from E_{II} in that the single mode of vibration is considered to be in its first excited state. It is then the gradual dropping out of the transition $E_{\text{II}}\rightleftarrows E_{\text{III}}$ from the effective heat capacity at high frequencies which produces dispersion and absorption. For the equation for adiabatic change (1.12) may then be written

$$-P\Delta v=(3/2)R\Delta T+(\partial(E_{\text{II}}-E_{\text{I}})/\partial T)_v\Delta T+(E_{\text{III}}-E_{\text{II}})\Delta n_1 \tag{2.1}$$

or, in more abbreviated form,

$$-P\Delta v=C_\infty\Delta T+E\Delta n_1. \tag{2.2}$$

This case will be discussed with pedantic attention to detail in order to illustrate the method of calculation.

In going from (2.1) to (2.2) a step has been taken which is not permissible without further justification. By including in C_∞, which represents the effective heat capacity of the gas at frequencies so high that $\Delta n_1=0$, all the heat capacity except that due to the $E_{\text{III}}\rightleftarrows E_{\text{II}}$ transition, two assumptions have been made. The first, which supposes that a terminal velocity corresponding to

$$V_\infty=\{(P/\rho)(1+R/C_\infty)\}^{\frac{1}{2}}$$

may actually be found can, at present, be justified only by experiment. The second, which supposes that, in case all energy states except one are adjusted at a rate much greater than the frequency, their influence on the acoustical behavior of the single slowly adjusted state is negligible, is justified by the more complicated models which are developed in Sections 3 and 4.[9a]

At mitall frequencies of sound the elasticity of the gas has equilibrium value, and (2.2) it becomes

$$-P\Delta v=C_\infty\Delta T+E(\partial n_1/\partial T)_v\Delta T=C_0\Delta T. \tag{2.3}$$

From Boltzmann's principle $n_1/n_0=e^{-x}$ and $n_0=(1+e^{-x})^{-1}$. Hence, in the three-state gas,

$$C_0-C_\infty=Rx^2e^{-x}/(1+e^{-x})^2. \tag{2.4}$$

It is now necessary to relate Δn_0 and Δn_1 to the acoustical frequency. This relation is given by *the reaction equation*, which may be written

$$-dn_0/dt=f_{01}n_0-f_{10}n_1 \tag{2.5}$$

if t denotes time and n_1, n_0 the fractions of molecules having the energy of the first excited state and the zero point state respectively. The

[9a] Bourgin (reference 9, 1932 paper) has demonstrated this explicity in his three-state model by carrying all rapidly adjusted states through the entire calculation with specific finite transition probabilities.

symbols f_{01} and f_{10} are used to denote the probabilities of the transitions $0\rightarrow1$ and $1\rightarrow0$ due to any cause; they are taken, for reasons which will later become apparent, as rates of excitation and de-excitation per molecule per second. Thus f_{10} is the mean number of times per second which a molecule in state 1 leaves state 1 for the state 0; it may be expected to depend on pressure and temperature. At equilibrium $f_{10}n_1 = f_{01}n_0$, and therefore $f_{01}=f_{10}e^{-x}$.

Eq. (2.5) is now displaced from equilibrium by small, finite increments of the independent variables, these increments being written Δn_0, ΔT, and ΔN. N as before is the numerical density and ΔT refers, as always, to the temperature of the translational terms. The three partial differential coefficients resulting from the displacement of the left-hand side will not be explicitly written. Since $\bar{n}_0$, the equilibrium value of n_0, has become $\bar{n}_0+\Delta n_0$ it is sufficient to write

$$d(\bar{n}_0+\Delta n_0)/dt = d\Delta n_0/dt = i\omega\Delta n_0, \quad (2.6a)$$

the evaluation of the time derivative following from the fact that Δn_0 varies about its amplitude according to a factor $e^{i\omega t}$. The displacement of the right-hand side involves three sets of partial coefficients. Displacement with respect to n_0 gives, if f_{01} and f_{10} are independent of n_0,

$$\left(\frac{\partial(f_{01}n_0-f_{10}n_1)}{\partial n_0}\right)_{TN}\Delta n_0 = -(f_{10}+f_{01})\Delta n_1, \quad (2.6b)$$

because in this case $\Delta n_0+\Delta n_1=0$. The displacement with respect to T gives, in close analogy to the van't Hoff isochore,

$$(\partial(f_{01}n_0-f_{10}n_1)/\partial T)_{n_0N}\Delta T = (f_{10}n_1x/T)\Delta T. \quad (2.6c)$$

The result of the change $N=\bar{N}+\Delta N$ may be represented in functional form, or some relation between f_{01}, f_{10} and pressure may be assumed. It is more convenient to take the latter course, and to assume that

$$f_{01}/N = k_{01} \quad \text{and} \quad f_{10}/N = k_{10}, \quad (2.7)$$

where k_{01} and k_{10} correspond to the bimolecular rate constants of the reaction $N_0+N\rightleftarrows N_1+N$. It is supposed, therefore, that the excitation of vibrational energy takes place solely because of a double collision of the first kind, and that its de-excitation is the result solely of a double collision of the second kind. It is further supposed that any molecule is equally capable of causing these transitions whether or not it is itself excited. This assumption can hardly be in strict accord with fact, and a divergence between the predicted and the experimental variation of the acoustical properties of the gas with pressure must therefore be expected. The interpretation of this discrepancy is touched on in Section 5. On the assumption (2.7) the pressure displacement yields

$$(\partial(f_{01}\Delta n_0-f_{10}\Delta n_1)/\partial N)_{Tn_0}\Delta N = (\Delta N/N)(f_0n_0-f_{10}n_1)=0. \quad (2.6d)$$

The displaced equation follows from (2.6a, b, c and d) and gives the result

$$E\Delta n_1 = Rx^2n_1\left(\frac{\Delta T}{1+e^{-x}+i\omega/f_{10}}\right) \quad (2.8)$$

and the constants for substitution in (1.14) and (1.15) are evidently $a=e^{-x}$, $b=0$, $c=(1+e^{-x})$, $d=\omega/f_{10}$, $n_0=(1+e^{-x})^{-1}$. The velocity of sound of small amplitude as a function of frequency will therefore be given by

$$V^2\mu = \frac{P}{\rho}\left\{1+R\frac{C_0(f_{10}+f_{01})^2+C_\infty\omega^2}{C_0^2(f_{10}+f_{01})^2+C_\infty^2\omega^2}\right\} \quad (2.9)$$

and this expression will also hold for the pressure dependence of the velocity insofar as (2.7) is valid. It is not yet suitable to describe the effect of temperature, since the relation between f_{10} and temperature has not been examined. The absorption of sound is correspondingly

$$h = 2\pi y[(Q^2-1)/(1+Q^2y^2)] \quad (2.10)$$

to the first order of small quantities if $Q=V_\infty/V_0$ and $y=C_\infty\omega/C_0(f_{10}+f_{01})$. These are essentially the expressions given by Kneser, Bourgin, and several others. Their properties and implications have been widely discussed.[11]

3. A Four-State Gas

The majority of gases which have been investigated do not conform to (2.9) within 0.1

[11] In addition to the references already given the papers of Luck, Phys. Rev. **40**, 440 (1932), and of Henry, Proc. Camb. Phil. Soc. **28**, 249 (1932), should be cited, since they deal with aspects of acoustical absorption.

percent. For carbon dioxide, the most popular gas, C_0 at 30°C is 3.435 R cal. deg.$^{-1}$, whereas $C_\infty+2Rx^2n_0[e^{-x}/(1+e^{-x})]$ is 3.291 R[12] and $V_{0\,\text{calc.}}$ is greater than $V_{0\,\text{obs.}}$ by about 0.5 percent. In these cases it is, of course, possible to use (2.9) as an approximation, but the meaning of $f_{01}+f_{10}$ is then obscured. More detailed analysis of π/Δ is needed to make available additional information. This may be exemplified by a model containing four internal energy states, E_{I}, E_{II} and E_{III} being as before, and E_{IV} being E_{III} plus the energy necessary to raise one gram molecule of the gas from the first to the second excited state of the slowly adjusted vibration.

The equation for adiabatic change is then

$$-P\Delta v=C_\infty\Delta T+E(\Delta n_1+2\Delta n_2) \qquad (3.1)$$

if the oscillator which represents the slowly adjusted vibration is harmonic (i.e., $E_{\text{III}}-E_{\text{II}}=E_{\text{IV}}-E_{\text{III}}=E$). Two reaction equations are now necessary

$$\left.\begin{aligned}-dn_0/dt&=(f_{01}+f_{02})n_0-f_{10}n_1-f_{20}n_2,\\ -dn_1/dt&=(f_{10}+f_{12})n_1-f_{01}n_0-f_{21}n_2.\end{aligned}\right\} \qquad (3.2)$$

Two assumptions will be made in displacing these from equilibrium. The first, which supposes $f_{01}=Nk_{01}$, $f_{12}=Nk_{12}$, etc., has been discussed at length in Section 2 under (2.7). The second supposes f_{02} and f_{20} to be much less than the other transition probabilities, and terms containing them consequently to be negligible. In the interpretation of experiment it must be remembered that these assumptions may not conform to fact. The displaced equations, together with the simultaneous equation $\Delta n_0+\Delta n_1+\Delta n_2=0$, give solutions for Δn_1 and Δn_2 in terms of the exponential factor x and the various transition probabilities. When Δn_1 and Δn_2 are substituted in (3.1) the appropriate form of the characteristic acoustical equation follows at once from (1.10). This has the form of (1.13) above where $n_0=(1+e^{-x}+e^{-2x})^{-1}$ and

$$\left.\begin{aligned}a&=e^{-x}+4e^{-2x}+e^{-3x},\\ b&=\omega(e^{-2x}/f_{10}+e^{-x}/f_{21}),\\ c&=1+e^{-x}+e^{-2x}-\omega^2/f_{10}f_{21},\\ d&=\omega(1+e^{-x})(1/f_{10}+1/f_{21}).\end{aligned}\right\} \qquad (3.3)$$

The velocity and absorption of sound follow from appropriate substitution in (1.14) and (1.15); the resulting equations are so complicated that a glance can hardly suffice to reveal their significance, and it appears therefore purposeless to print them in full. A few of the reductions of the expression for the velocity may, however, be examined with profit. For example, when $\omega\ll f_{10}, f_{21}$

$$V_0^2=\frac{P}{\rho}\left\{1+\frac{R}{C_\infty+Rx^2n_0[(e^{-x}+4e^{-2x}+e^{-3x})/(1+e^{-x}+e^{-2x})]}\right\}, \qquad (3.4)$$

which represents the terminal velocity at low frequency. When $\omega\gg f_{10}, f_{21}$

$$V_\infty^2=(P/\rho)\{1+R/C_\infty\} \qquad (3.5)$$

gives the terminal velocity at the highest frequency for which the model is adapted. This is of course not infinite; the terms in C_∞ are still considered in equilibrium in all parts of the wave. When $f_{21}\ll\omega\ll f_{10}$ an intermediate constant velocity

$$V^2_{f_{21}\doteq 0}=\frac{P}{\rho}\left\{1+\frac{R}{C_\infty+Rx^2n_0(e^{-x}/(1+e^{-x}))}\right\} \qquad (3.6)$$

is reached, and when $f_{10}\ll\omega\ll f_{21}$

$$V^2_{f_{10}\doteq 0}=\frac{P}{\rho}\left\{1+\frac{R}{C_\infty+Rx^2n_0(e^{-2x}/(1+e^{-x}))}\right\}, \qquad (3.7)$$

the alternate intermediate constant velocity, is obtained. A third intermediate constant velocity, due to the complete failure of transitions to and from the 1st state, may of course be imagined, but is considered so unlikely to be realized in practice that it has not been provided for in the four-state model above. It may simply be

[12] Both these values have been calculated on the basis of fundamental frequencies 667.5, 1285, and 2326 cm^{-1}, as reported by Villars, Chem. Revs. 11, 408 (1932). C_∞ is taken as 2.5085R cal. deg.$^{-1}$ in the second.

evaluated by taking

$$\partial n_2/\partial T = -\partial(n_0 e^{-2x})/\partial T$$

which gives

$$V_{f_{10}\doteq f_{12}\doteq 0} = \frac{P}{\rho}\left\{1+\frac{R}{C_\infty + 4Rx_2 n_0(e^{-2x}/(1+e^{-2x}))}\right\}. \quad (3.8)$$

The practical realization of constant velocities corresponding to (3.6), (3.7) and (3.8) does not appear very probable, but the influence of a decided difference between f_{10} and f_{21} may be detected by experiment.

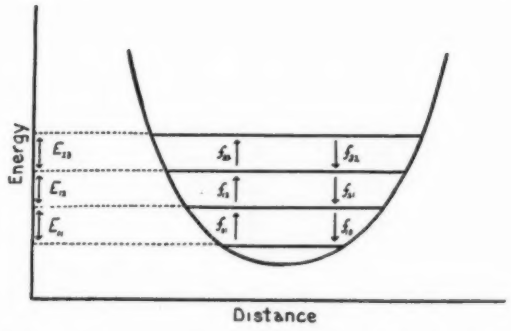

FIG. 1. Diagrammatic representation of the frequency-dependent part of the heat capacity for the five-state model. Transitions involving energy changes greater than $\pm h\nu_0$ have been considered of a low order of probability, and harmonicity of the oscillator has been postulated by setting $E_{01}=E_{12}=E_{23}=h\nu_0$. Where these approximations are permissible the model may be applied to any mode of vibration, whether transverse or longitudinal.

4. A Five-State Gas

Although the four-state gas model developed in the previous section is adequate to describe the behavior of carbon dioxide at all ordinary temperatures, it cannot be used for gases in which the fundamental vibrational frequency is much lower. It is exactly in these gases, however, that differences in the transition probabilities of the various states may most easily be investigated. It is advisable, therefore, to carry the development of π/Δ one step further, and examine a five-state model. As before the energy states $E_{II}-E_I$ will be considered in equilibrium with the adiabatic cycle at all frequencies. The frequency-variable transitions have been illustrated in Fig. 1, which, with the explanations in the previous sections, gives the model sufficient definition. The development has been carried out exactly as in the previous section with the help of the assumptions that the probabilities are directly proportional to pressure, that all transitions involving an energy change greater than $\pm h\nu$ are so rare as to be negligible, and that the oscillator is harmonic. Three appropriate reaction equations are displaced from equilibrium, and yield, with the equation $\Delta n_0+\Delta n_1+\Delta n_2+\Delta n_3=0$

$$\left.\begin{aligned}
\Delta n_0 &= \frac{\Delta n_1 e^x - x n_0 \Delta T/T}{i\omega e^x/f_{10}+1}\\
\Delta n_1 &= (i\omega/f_{21}+1)\Delta n_2 e^x + (i\omega/f_{21})\Delta n_3 e^x - x n_0 e^{-x}(\Delta T/T)\\
\Delta n_2 &= (i\omega/f_{32}+1)\Delta n_3 e^x - x n_0 e^{-2x}(\Delta T/T)
\end{aligned}\right\}. \quad (4.1)$$

These are solved simultaneously and substituted in the equation for adiabatic change

$$-P\Delta v = C_\infty \Delta T + E(\Delta n_1 + 2\Delta n_2 + 3\Delta n_3) \quad (4.2)$$

to evaluate the constants for substitution in (1.14) and (1.15). These are $n_0=(1+e^{-x}+e^{-2x}+e^{-3x})^{-1}$

$$\left.\begin{aligned}
a &= e^{-x}+4e^{-2x}+10e^{-3x}+4e^{-4x}+e^{-5x}-\omega^2(e^{-3x}/f_{10}f_{21}+e^{-2x}/f_{10}f_{32}+e^{-x}/f_{21}f_{32})\\
b &= \omega[(e^{-x}+4e^{-2x}+e^{-3x})(e^{-x}/f_{10}+1/f_{32})+(e^{-x}+e^{-2x}+e^{-3x}+e^{-4x})/f_{21}]\\
c &= 1+e^{-x}+e^{-2x}+e^{-3x}-\omega^2(1+e^{-x})(1/f_{10}f_{21}+1/f_{10}f_{32}+1/f_{21}f_{32})\\
d &= \omega(1+e^{-x}+e^{-2x})(1/f_{10}+1/f_{21}+1/f_{32})+\omega e^{-x}/f_{21}-\omega^3/f_{10}f_{21}f_{32}
\end{aligned}\right\}. \quad (4.3)$$

Only two of the reductions of the resulting expression for the velocity need be examined. When

$\omega \ll f_{10}, f_{21}, f_{32}$

$$V_0^2 = \frac{P}{\rho}\left\{1+\frac{R}{C_\infty + Rx^2[(e^{-x}+4e^{-2x}+10e^{-3x}+4e^{-4x}+e^{-5x})/(1+e^{-x}+e^{-2x}+e^{-3x})^2]}\right\} \tag{4.4}$$

and the low frequency terminal velocity is obtained. The high frequency terminal velocity is identical in form with (3.5). If $f_{32} \ll \omega \ll f_{10}, f_{21}$

$$V^2_{f_{32}=0} = \frac{P}{\rho}\left\{1+\frac{R}{C_\infty + Rx^2 n_0[(e^{-x}+4e^{-2x}+e^{-3x})/(1+e^{-x}+e^{-2x})]}\right\}, \tag{4.5}$$

which represents a wholly different situation from (3.4) although the two expressions bear a strong superficial resemblance to each other. The other intermediate constant velocities are easily obtained by suitable modifications of the constants given in (4.3). They will not be printed.

5. The Behavior of a Gas with Many Independently Adjusted Energy States

Before proceeding further it will be well to discuss the experimental bearing of what has already been written. This will be undertaken from two viewpoints: first, to discover what fresh information has become available and second, to find reasonably satisfactory approximations which obviate, at any rate during the course of an experimental research, the necessity of calculating from the cumbersome equations resulting from (1.14) and (1.15) with appropriate constants. Since a relation between absorption and velocity is already at hand in (1.19), the velocity of sound will alone be given attention below.

Figs. 2 and 3 illustrate the pressure and frequency dependence of the models developed in Sections 3 and 4, respectively. In each case the ordinates are expressed as the "apparent heat capacity ratio" which is an abbreviation for the apparent ratio of heat capacity at constant pressure to that at constant volume at a particular value of ω. The abscissae scale corresponding values of ω/f_{01} which have, however, been expressed as ω/Nk_{01} since the foregoing exposition is committed to the formal equality of these quantities. This method of plotting describes, therefore, with equal suitability, the behavior of the models at constant frequency and variable pressure, or at constant pressure and variable frequency. Fig. 1 has been obtained by calculating numerically, from the constants (3.3) the cases $f_{10}=f_{21}$, $10f_{10}=f_{21}$, and $100f_{10}=f_{21}$ in a gas for which the exponent $h\nu_0/kT$ is 3.00. Fig. 3 gives numerical values for the effective heat capacity ratio according to (4.3) in a gas where $h\nu_0/kT$ is 2.00 for the cases $f_{10}=f_{21}=f_{32}$, $100f_{10}=10f_{21}=f_{32}$ and $10^4 f_{10}=10^2 f_{21}=f_{32}$. To avoid ambiguity the slowly adjusted vibration has been considered to have an *a priori* probability of unity, although this will not be the case when

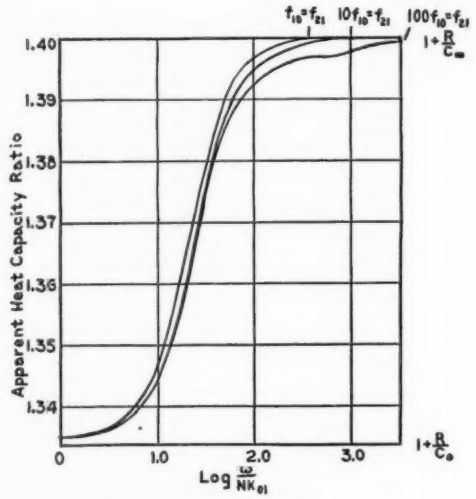

FIG. 2. The acoustical behavior of a model in which the frequency-dependent part of the heat capacity is adequately described by the $0 \rightleftarrows 1$ and $1 \rightleftarrows 2$ transitions. The ordinates represent the apparent values of C_p/C_v and consequently indicate also the variations of V^2. The abscissae give corresponding values for log ω/f_{01}, and serve therefore to describe the behavior of the model at constant pressure and variable frequency or, conversely, at constant frequency and variable pressure if the assumption $f_{01}=Nk_{01}$, which has been discussed in the text, is valid. The value of $h\nu_0/kT$ has been chosen as 3.0 for the purpose of plotting, and it has been assumed that the frequency-dependent part of the heat capacity is due to a single valence vibration, the contribution of all other vibrations being negligible.

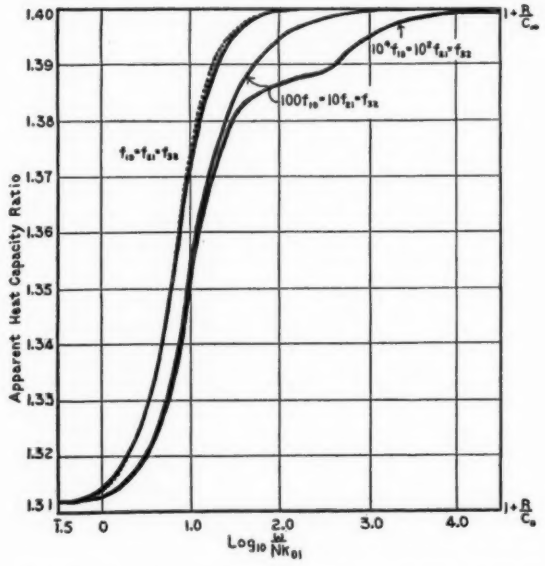

FIG. 3. The acoustical behavior of a model in which the frequency-dependent part of the heat capacity is adequately described by the $0 \rightleftarrows 1$, $1 \rightleftarrows 2$, and $2 \rightleftarrows 3$ transitions of a valence vibration. In the case plotted $h\nu_0/kT$ has been given a value of 2. (This is suitable for comparison with carbon bisulfide at ordinary temperatures if the value of $C_0 - C_\infty$ is doubled throughout.) The graph indicates that differences in the probabilities f_{01} and f_{12} of one order of magnitude should be detectable by measurements at constant frequency and variable pressure or at constant pressure and variable frequency.

the dispersive region is caused by the failure at high frequencies of a degenerate vibration. C_∞ has been uniformly assigned the arbitrary value $2.5R$.

The dependence of the velocity of sound on frequency at constant pressure and temperature

Several conclusions follow from the consideration of Figs. 2 and 3.

I. If the transitions $1 \rightarrow 0$, $2 \rightarrow 1$, $3 \rightarrow 2 \cdots$ have the same probabilities, the square of the measured velocity of sound plotted as a function of the logarithm of the frequency will conform within experimental accuracy to a symmetrical *S*-shaped curve. An arbitrary number of energy states cannot of course be treated by the method used above, but there appears to be no danger in making this statement general.

II. Any detectable deviation from a symmetrical *S*-shaped curve may be attributed to a difference in the transition probabilities f_{10}, f_{21} $\cdots$. Such a difference has been predicted by Zener,[13] and its direct experimental demonstration should not be difficult by this means. If the higher energy states are more rapidly adjusted than the lower, the high frequency part of the *S*-curve is flattened; if they are less rapidly adjusted than the lower, the curve is flattened in its low frequency region.

III. If the difference between the rates of adjustment of any two states is greater than about two orders of magnitude, the deformed *S*-curve becomes irregular, and two or more maxima appear in the variation of the absorption coefficient with frequency. The appearance and separation of these maxima are conditioned by the relative importance of the two energy states in the heat capacity.

IV. If the rate of adjustment of any state is greater than the others by about three orders of magnitude, its contribution to the heat capacity may, from the point of view of experiment, legitimately be incorporated in C_∞ and not assigned a specific rate of adjustment. This is the justification for giving $\partial(E_{II} - E_I)/\partial T$ its equilibrium value throughout the dispersive region in the proceeding sections.

V. If the transitions $2 \rightarrow 0$, $3 \rightarrow 0$, $3 \rightarrow 1$, etc., involving energy changes of two or more quanta are not less frequent than those $1 \rightarrow 0$, $2 \rightarrow 1$, etc., by about two orders of magnitude the low frequency portion of the *S*-curve may be deformed, and the only legitimate remedy is to describe the situation by a model containing a greater number of transitions. Since, however, the maximum possible contribution to the heat capacity of these transitions is small, their disturbing effect will probably escape experimental detection. The uncongenial task of still further complicating the calculations in this way will be postponed in the fervent hope that experiment may not demand it.

The dependence of the velocity of sound on pressure at constant frequency and temperature

A purely formal consideration of the relationships given in Sections 2, 3 and 4 leads to the conclusion that the increase of the frequency by a certain factor should be exactly balanced by

[13] Zener, Phys. Rev. **37**, 556 (1931); **38**, 277 (1931).

the decrease of the pressure by the same factor. All the conclusions just stated for the dependence of the velocity of sound on frequency therefore appear to hold in the inverse sense for its dependence on pressure. The correspondence between frequency and pressure is due, however, to the simplifying assumptions and approximations which have been used. In practice the conclusions reached from considering frequency may be applied to pressure only with the following important qualifications.

I. In a real gas the heat capacity at constant volume is not independent of pressure, and pressure and numerical density are not exactly proportional. Although it is always possible to eliminate P/ρ by dividing V_ω by V_0 at any given pressure or frequency, errors due to these causes may become experimentally manifest at high pressures.

II. The assumption that f_{10}, f_{21}, etc., vary directly in proportion to the pressure may be invalid on several grounds. If, for some reason which is not at present clear, triple collisions are more effective than double collisions in adjusting the internal energy $f_{10} \neq Nk_{10}$. The mathematical treatment will not seriously be disturbed by this since (2.6d) will vanish for any reasonable functional relationship between f_{10}, etc., and N. A diminution of the pressure will, however, have more effect on the velocity and absorption than a corresponding increase of frequency. If the advantage of triple collisions is marked, this effect should be very strikingly manifest. If only slight, it may escape experimental detection because of the relative infrequency of triple collisions. A correction, if necessary, may be applied to the high-pressure region by expressing each f as the sum of two probabilities, one of which is proportional to the number of double collisions, and one to the number of triple collisions.

III. Again, it is possible that a molecule which is vibrating has an advantage over one which is not vibrating in bringing about energy transformations. Since no example of this is yet known, we need not concern ourselves with a mechanism for it, but merely with its effect on the acoustical behavior of the gas. A more complete statement of the assumption $f_{10}=k_{10}N$ reads $f_{10}=k_{10}Nn_0+k'_{10}Nn_1+k''_{10}Nn_2+\cdots$ where $k_{10}\equiv k'_{10}\equiv k''_{10}\cdots$. If this identity is for any reason invalid, $\partial f_{10}/\partial n_1$ no longer vanishes, and (2.6b) is incorrect. A variation in n_1 may, in a perfect gas, be brought about only by changing the temperature, and this introduces a set of approximations which have not yet been considered. Both for this reason, and because n_1, n_2, etc., are usually small fractions, it seems unlikely that any inequality of k_{10}, k'_{10}, etc., may at present be demonstrated by experiment in pure gases.

The presence of another mode of vibration which, although excited to an appreciable extent, remains in equilibrium with the sound wave throughout the dispersive region, may, of course, cause exactly similar effects. There is some reason to believe, however, that resonance between the two vibrational modes must be close in order for collisions of this type to be effective.

IV. When the dispersive region is produced by the failure of transitions between vibrational states which are optically active, the excited molecules may radiate their energy between collisions. The effect of this will be to cause the mean lifetime of the molecules to increase less rapidly with decreasing pressure than is indicated by the purely kinetic assumption $f_{10}=Nk_{10}$. This possibility was first envisaged by Bourgin,[14] who has pointed out that the effect becomes important at low pressures. It has recently been stressed, especially in connection with acoustical absorption, by Kneser.[15] It is not intended to extend the scope of this study to cover radiation effects. It will therefore merely be pointed out that if the relaxation time is written as $\vartheta=1/[N(k_{10}+k_{01})+1/\tau]$ where τ is a temperature and pressure-independent quantity expressing the radiation lifetime, a rough estimate of τ may be obtained from measurements of the velocity of sound as a function of pressure in the dispersive region. Since the temperature and pressure derivatives of τ vanish, its introduction in this way does not disturb the exposition given in Sections 2, 3, and 4. The radiation lifetimes of molecules are probably so great that it is not to be expected that radiation effects will become appreciable at pressures which are at present suitable for the measurement of the

[14] Bourgin, reference 9, 1928 papers.
[15] Kneser, reference 6, especially 1933 paper.

velocity of sound. It should be noted, however, that acoustical measurements at very low pressures may ultimately become an effective means of investigating radiation lifetimes.

The dependence of the velocity of sound on temperature at constant frequency and numerical density

The numerical density of a gas increases with decreasing temperature at constant pressure, and it will therefore be convenient to discuss the variation of the velocity of sound with temperature at constant numerical density rather than at constant pressure. Some assumption must now be made concerning the temperature dependence of C_∞ and the f probabilities.

C_∞ will depend on temperature if some of the non-lagging degrees of freedom are not fully excited.[16] In this case $C_\infty=\frac{3}{2}R+p_1Rx_1^2e^{-x_1}/(1-e^{-x_1})^2+\cdots$ if p_1, etc., are the *a priori* probabilities of the degrees of freedom, and x_1, etc., abbreviations for $h\nu_1/kT$, etc.

In discussing the temperature coefficient of the f probabilities only those of the type f_{10}, f_{21}, etc., will be considered, since the relation between these and the f_{01}, f_{12}, etc., probabilities is already known. The simplest assumption is the C_∞ and f_{10}, etc., are independent of temperature. In this case the three, four, and five-energy state models behave similarly: the S-curve is merely compressed as the temperature is lowered. This is roughly described by stating that $(V_\omega^2-V_\infty^2)/(V_0^2-V_\infty^2)$ is independent of temperature for a given value of ω. Fig. 4 diagrammatically illustrates this situation.

Such measurements as exist on the temperature dependence[17] of the velocity of sound in the dispersive region indicate that the transition probabilities in several gases are by no means independent of temperature.

To account for this a quantity called the

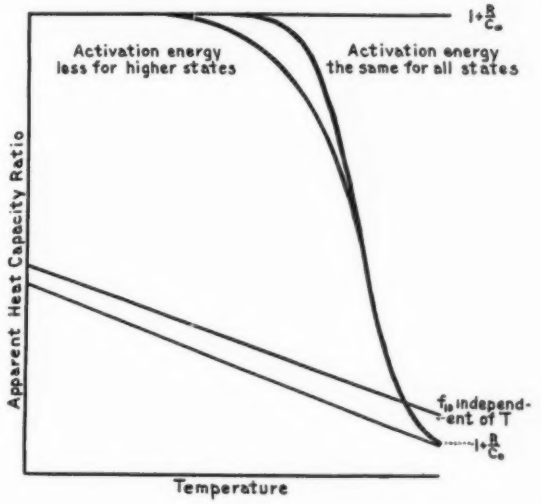

FIG. 4. Diagrammatic representation of various possibilities for the temperature dependence of the velocity of sound at constant frequency and constant numerical density. In practice f_{10} will depend on the number of collisions per unit time, and the straight line "f_{10} independent of T" is therefore an idealization. Relatively small differences between the activation energies of collision of two states should be experimentally manifest by measuring the temperature coefficient of the velocity at the low frequency and high frequency limits of the dispersive region.

activation energy of collision has already been suggested.[17] This bears an analogy to the activation energy of a chemical reaction in that

$$\frac{\partial f_{10}}{\partial T}=\frac{f_{10}A_{10}}{RT^2},\quad \frac{\partial f_{01}}{\partial T}=\frac{f_{01}(A_{10}+E)}{RT^2},\quad \text{etc.}$$

and represents the "threshold kinetic energy" necessary to effect the transitions $1\rightarrow 0$, etc. Such an impediment to the transference of vibrational to kinetic energy might have several causes. For example Heil,[18] following Oldenberg,[19] has suggested that the dynamics of collisions of the second kind demand that in some cases only molecules of particularly high kinetic energy should be effective. On the other hand, a potential energy barrier of the same general type that found in chemical reactions might be imagined.[17, 20] This possibility will be discussed at slightly greater length in order to make clear

[16] Kneser and Zühlke, Zeits. f. Physik **77**, 649 (1932), have shown in nitrous oxide that the deformation vibration fails at frequencies lower than the other vibrations. Similar evidence concerning carbon disulphide has recently been obtained in the writers' laboratory in Princeton by Mr. J. A. Reid.

[17] Apparently the measurements of Richards and Reid, Nature **130**, 739 (1932), are the only ones reported present which deal directly with this question.

[18] Heil, Zeits. f. Physik **14**, 31 (1932).

[19] Oldenberg, Phys. Rev. **37**, 194 (1931).

[20] Franck and Eucken, Zeits. f. physik. Chemie **A20**, 460 (1933).

certain characteristics which the activation energy of collision may conceivably possess.

If two molecules approach one another with a given relative kinetic energy they influence increasingly each other's chemical binding energies as they come increasingly into each other's repulsive fields. This process will, with a certain probability, induce a change in the vibrational states of the colliding molecules. A simple though approximate method of calculating the transition probability has been given by Kallmann and London.[21] It may be assumed, in analogy to their mechanism, that the process of approach of the two molecules occupies an extremely short period of time, and that during this time the vibrational wave function remains unchanged. This means that, owing to the mutual "softening" of the chemical bonds,[21a] there is a certain probability that a change will occur in the vibrational quantum numbers of one or both of the colliding molecules. The process of separation is, conversely, considered to be very slow, so that no further change of the vibrational quantum numbers occurs; this must be assumed in order to have a nonvanishing transition probability, since in a rapid separation the wave function would again remain unchanged, and the vibrational quantum numbers would revert to their original values.

During the second, or adiabatic, process the hardening of the bonds as the molecules separate will effect an increase of the vibrational energy. If ν represents the frequency of the original bond, and ν' that of the softened bond, the initial vibrational energy may be increased by a factor, ν/ν'. For a molecule originally in the nth state this excess of energy, which is supplied at the expense of the relative translational energy, will amount to $(\nu/\nu')(n+1/2)h\nu-(n+1/2)h\nu$, and the probability of an excitation will be roughly in proportion to this divided by $h\nu$, or to $(\nu-\nu')/\nu(n+1/2)$. This quantity increases rapidly with increasing values of n, and is three times larger for the second than for the first vibrational state. Thus when n is small $(\nu-\nu')/\nu$ must be relatively larger, and a greater softening of the bond is necessary to effect an increase of vibrational energy. This can be brought about only by closer approach of the molecules, and consequently by collisions involving greater relative kinetic energies. The difference will be experimentally manifest by larger activation energies of collision for the excitation of the lower than of the higher quantum states.

While the mechanism described above is incorrect in that the rapidity of approach and separation must, in actuality, be of the same order of magnitude, it is believed to represent a plausible picture for the expected relationship $A_{10}>A_{21}>A_{32}\ldots$. It is hoped that a more satisfactory description of energy transitions during collision of the first and second kinds may be obtained from a quite different method of approach, which cannot here be included.

It is not proposed, however, to introduce the activation energy of collision on deductive grounds, but to use it as a purely empirical quantity calculable from the temperature coefficients of f_{10}, f_{21}, etc., according to the expressions derived from (1.14) and (1.15) in appropriate cases. Any reasoning from analogy with the chemical activation energy will be suppressed, since there is at present no convincing reason for believing the activation energy of collision to be independent of temperature over any considerable range. Furthermore, it is not at present possible to predict its order of magnitude. It has been introduced only to facilitate the description of the temperature coefficient of the velocity in terms of experimentally determined quantities, and to permit a rough kinetic analysis of the probabilities f_{10}, f_{21}, etc.

Before proceeding further it must be pointed out that the existence of an activation energy of collision introduces a new approximation into the treatment of the previous sections. If the molecules on the high energy side of the Maxwell-Boltzmann distribution lose translational energy to vibrational energy on every collision, the effective pressure of the gas must also vary somewhat with frequency in the dispersive region. It does not appear possible at present to compute the extent of this approximation, but it is believed to be of small importance.

[21] Kallmann and London, Zeits. f. physik. Chemie **B2**, 207 (1929).

[21a] The "softening" of the chemical bonds as a forerunner of chemical decomposition has been studied, in connection with Raman spectra, by Cremer and Polanyi, Zeits. f. physik. Chemie, Bodenstein-Festband, 771 (1931).

The effect of an activation energy of collision on the dispersive behavior of the gas is clear. If f_{10}, f_{21}, etc., all have equal activation energies, lowering the temperature will correspond to increasing the frequency, and the gas may be made to pass through its dispersive region by means of temperature changes alone. This case has been diagrammatically represented in Fig. 4; the S-curve is slightly deformed by the variation of C_0 with temperature. The greater the activation energy the more rapidly will the gas pass through its dispersive region as the temperature is lowered. If the activation energies are different for f_{10}, f_{21}, etc., the S-curve will be more greatly deformed. Since it seems probable, on the argument given above, that $A_{10}>A_{21}$, etc., this case has also been diagrammatically represented in Fig. 4. The extreme upper and lower temperature of the dispersive region will in this case show different activation energies of collision, which may be interpreted by means of a gas model containing a suitable number of energy states.

The kinetic analysis of transition probabilities

All modern writers on the dispersion and absorption of sound have used a kinetic mechanism as the basis for their interpretation.[22] If experiment has established the magnitude of the activation energy of collision and the essential importance of double collisions in the excitation process, a little further information may be obtained by subjecting the transition probabilities to a rough kinetic analysis.[17] The results so obtained must not, however, be considered significant in any exact quantitative sense.

The number of double collisions suffered by the molecules in the 0 state per cubic centimeter per second is approximately

$$Z_0=4N^2n_0\sigma^2(\pi RT/M)^{\frac{1}{2}}, \qquad (5.1)$$

where N is the total number cc^{-1} and n_0 the fraction of these in the 0 state as before, M is the molecular weight, and σ the distance of nearest approach of the molecules in a double collision. On the assumption that σ is unaltered by the changes of state $0\rightarrow1$, $0\rightarrow2$, etc., similar expressions for Z_1, Z_2, etc., follow. Carrying through the previous assumption that either steric or energetic properties of the molecules may hinder the transformation of kinetic into vibrational energy, the number of collisions cc^{-1} sec.$^{-1}$ which are effective in removing a molecule from the 0 to the 1 vibrational state may be written.

$$z_0=4N_2n_0\xi_{01}\sigma^2(\pi RT/M)^{\frac{1}{2}}e^{-A_{01}/RT}, \qquad (5.2)$$

where $2\pi\xi_{01}\sigma^2$ is a steric factor not strongly dependent on temperature, which may be termed the "effective cross section" of the molecule for the transition $0\rightarrow1$. Since the f probabilities are expressed per molecule per second it is apparent that to find f_{01} z_0 must be divided by Nn_0, the number of molecules in the 0 state cc^{-1}. Thus

$$f_{01}=(4N\sigma^2(\pi RT/M)^{\frac{1}{2}})\xi_{01}e^{-(A_{10}+E)/RT} \qquad (5.3)$$

and

$$f_{10}=(4N\sigma^2(\pi RT/M)^{\frac{1}{2}})\xi_{10}e^{-A_{10}/RT}, \qquad (5.4)$$

from which the equality of the effective cross sections in the two directions is immediately deducible from $f_{01}/f_{10}=e^{-x}$. Similar relations between any pair of transition probabilities $f_{n,\,n+1}f_{n+1,\,n}$ may similarly be written. There is no reason at present for supposing the various ξ values to be equal for different states.

From the comparison of $2\pi\sigma^2\xi_{01}$ with the ordinary kinetic cross section deduced from viscosity, effusion, and the like, it is possible that interesting information concerning the spatial characteristics of the excitation process may be deduced. For example if a triatomic colinear molecule is visualized by the naif model of three balls connected by springs, a "head on" collision of another molecule with either of the end atoms must be far less effective in exciting the deformation vibration than a "broadside" collision of another molecule with the center atom. The effective solid angle for collisions may in this case be small, and the ordinary cross section far greater than that indicated by acoustical measurements.[23]

It should be noted that if the exponential

[22] Apparently Jeans (*Dynamical Theory of Gases*, Cambridge (1904)), was the first to take this standpoint. Einstein, Herzfeld and Rice, Bourgin, and others have suggested or stressed a kinetic mechanism.

[23] Heil, reference 18, has shown this from Kneser's measurements on carbon dioxide.

factor in (5.4) is near to unity the effective cross section must be expressed with greater circumspection than has here been considered necessary, since in this case it may dominate the temperature coefficient of the dispersion. The probabilities f_{10}, etc., must also be expected to diminish slightly with temperature owing to the decreased frequency of collisions.

Approximations suitable for orientation in experimental investigations

Since calculation from the four and five-state modifications of (1.14) and (1.15) is a laborious and uncongenial process, two approximations are here given which it is hoped will be useful in orienting the experimenter during the course of his measurements. They are in certain cases also suitable for the final expression of measurements.

To describe the frequency and pressure dependence of the velocity the expression

$$V^2\mu=\frac{P}{\rho}\left\{1+\frac{RC_0+RC_\infty\omega^2\vartheta_N{}^2}{C_0{}^2+C_\infty{}^2\omega^2\vartheta_N{}^2}\right\}, \qquad (5.5)$$

where $\vartheta_N=(e^x+1)/(e^x-1)^2f_{01}$, is very nearly identical in its behavior to that of the multistate models of Sections 3 and 4 when $f_{10}=f_{21}=f_{32}$. The correspondence with the four state model is so close that no difference is apparent in Fig. 2. The correspondence with the five state model is within all reasonable experimental limits, as the dotted line in Fig. 3 shows. Any experimentally detectable deviation from (5.5) may therefore be attributed to inequality of f_{10}, f_{21}, etc., if the frequency dependence of the velocity is in question. Conversely if f_{10}, f_{21}, etc., have been proved identical by varying the frequency at constant pressure and temperature, deviation from (5.5) may be attributed to radiation or triple collision effects if the pressure dependence of the velocity is receiving attention.

It has been proposed to study the temperature coefficient of the velocity in the dispersive region merely to determine the activation energy of collision for each of the various excitation processes. Since this is given by f_{10}/f'_{10}, etc., corresponding to T and T', an error in the absolute magnitude of f_{10}, etc., is of no importance. It is somewhat better for the determination of the activation energy of collision to use the approximation

$$V^2\mu=\frac{P}{\rho}\left\{1+\frac{RC_0+RC_\infty\omega^2\vartheta_T{}^2}{C_0{}^2+C_\infty{}^2\omega^2\vartheta_T{}^2}\right\}, \qquad (5.6)$$

where $\vartheta_T=1/f_{10}(1+e^{-x})$ rather than the approximation (5.5). The values of f_{10} so obtained may be noticeably in error, but this will not affect the calculated activation energy. If A_{10} is different from A_{21}, etc., the low and high frequency parts of the S-curve will show different temperature dependence, as illustrated by Fig. 4. The exact significance of the difference can be demonstrated only by a more profound analysis, but its existence may conveniently be demonstrated by (5.6).

Both (5.5) and (5.6) assume a knowledge of the energy levels of the molecule, since they require the value of e^x. Cases may occur where these are not obtainable. A rough analysis may then be made by supposing C_0-C_∞[24] to be due to the failure of a single vibrational transition involving a suitable quantum jump. This will, of course, lead to more or less erroneous f values, but will give a variation of ϑ_T with temperature which is approximately correct.

6. A Few Further Cases. Gaseous Mixtures

It is believed that the suggestions given in the previous sections are sufficient to interpret the acoustical behavior of all pure gases which have at present been examined experimentally. They by no means suffice, however, to cover all possible cases of interest. These are, indeed, so numerous and diverse that the labor and space required to examine them individually are prohibitive. Since the description of a system of any necessary complexity by the method of Einstein or of Bourgin is a matter of labor rather than of ingenuity, more complicated cases may reasonably be postponed until required by experiment. The purpose of this section is only to give a superficial survey of further territory.

The simultaneous failure of two modes of vibration

Kneser's contention that in gases in which the deformation vibration is excited it is this vibra-

[24] This must in any case be known in order to identify the energy transitions which are absent at high frequencies.

tion alone which causes the first dispersive region[25] appears to be justified in all cases which have at present been reported. It has nevertheless been noted that chlorine and oxygen, which can have only longitudinal vibrations, show dispersion of sound in the region of frequency which is at present accessible to experiment, and the simultaneous failure of two vibrational modes in the acoustical cycle must therefore be considered as possible. If the excitation and de-excitation of the two modes is due only to the relative kinetic energy of the two molecules on collision, the behavior of the gas will be a composite of two examples of the type already discussed. An approximate description is obtained by writing for the characteristic acoustical equation

$$\frac{\pi}{\Delta}=\frac{P}{\rho}\left\{1+\frac{R(1+i\omega\vartheta_1)(1+i\omega\vartheta_2)}{C_\infty(1+i\omega\vartheta_1)(1+i\omega\vartheta_2)+C_1(1+i\omega\vartheta_2)+C_2(1+i\omega\vartheta_1)}\right\}, \tag{6.1}$$

in which, of course, $C_0=C_\infty+C_1+C_2$ and ϑ_1 and ϑ_2 are the approximate relaxation times of the vibrational modes k and l, respectively. Serviceable if somewhat inexact expressions for velocity and absorption follow from (6.1) by (1.6) and (1.7). The absorption will show maxima at two frequencies if ϑ_k and ϑ_l differ by about two orders of magnitude.

A more interesting case is provided if vibrational resonance phenomena increase the probability of the exchange of energy between the two modes of vibration. Such a possibility has already been touched on in Section 5, paragraph III of the discussion dealing with the variation of the velocity of sound with pressure. Each transition may then be brought about by one of several means, and it will in general be the most effective of these which determines the position of the dispersive region. The probability f_{01} will, for example, be the sum of the number of times per second which a molecule leaves the 0 state for the first state due to each of the various types of collisions. It is clear that unless a considerable fraction of the molecules are excited the number of transitions per second due to collisions with vibrating molecules will escape detection. It is also clear that, since the rate of transition due to each type of collision cannot be measured separately, a large number of extremely accurate measurements at different temperatures and pressures must be made in order to separate one from the other. It is therefore more likely that information of this type will be obtained from the study of gaseous mixtures than from pure gases.

[25] Kneser, reference 6, 1933 paper.

The behavior of a dissociating gas in which the heat capacity may vary with frequency has been treated at great length in Parts I and Ia of this series. It is perhaps worthy of note that if the conclusions stated above are justified this treatment is adequate for all present purposes.

Binary gaseous mixtures

The discussion of the previous paragraphs immediately suggests a parallel treatment for gaseous mixtures. This branch of acoustical theory has already been comprehensively developed by Bourgin and nothing of importance need here be added to what he has said. The following remarks are primarily to be interpreted as a translation of Bourgin's work into the language of Einstein's theory. The analysis of the transition probabilities is, however, somewhat more specific, and consequently more vulnerable than that which Bourgin has provided, although it is believed to be fully in accord with his intentions.

The simplest dispersive binary gaseous mixture consists of a dispersive gas A and a nondispersive gas B. The heat capacity in equilibrium of A is C_{0A} mole^{-1} deg.$^{-1}$, and of this a part $C_{0A}-C_{\infty A}$ is dependent on frequency. The heat capacity of B is C_{0B} at all frequencies which are experimentally considered. The numerical density of A is mN and of $B(1-m)N$. The effective heat capacity at very low frequencies of sound will then be $C_0=C_{0A}m+C_{0B}(1-m)$ and at high frequencies will be $C_\infty=C_{\infty A}m+C_{0B}(1-m)$. It will probably be sufficient to describe the acoustical behavior of A by a three-state approximation, particularly if measurements are compared at a single frequency in the dispersive region. In this

case the characteristic acoustical equation of the mixture is

$$\frac{\pi}{\Delta}=\frac{P}{\rho_A m+\rho_B(1-m)}\left\{1+\frac{R(1+i\omega\vartheta)}{C_\infty(1+i\omega\vartheta)+(C_{0A}-C_{\infty A})m}\right\} \tag{6.2}$$

and expressions for the velocity and absorption follow from (1.6) and (1.7). P in (6.2) is of course the sum of the partial pressures of A and B. The relaxation time ϑ is now, however, a more complicated quantity; collisions of an A molecule with another of the same species or with a B molecule may produce energy transitions, and it is the more effective of these encounters which will dominate the relaxation-time. This may be expressed by writing, for a pressure great enough to exclude radiation effects,

$$\vartheta \leqq \text{const.}/(f_{AA}+f_{AB})(1+e^{-x}), \tag{6.3}$$

where f_{AA} and f_{AB} are the probabilities per molecule per second of removing vibrational energy from an excited state by collisions of the type AA and AB, respectively. The constant in (6.3) may be taken as unity without losing the significance of the measurements; it will become larger than this if $C_{0A}-C_{\infty A}$ is great. The values of f_{AA} are obtained from measurements on pure A at various temperatures and partial pressures. In this way the activation energy of collision and effective cross section for collisions of the type AA may be evaluated by means of (5.3). Measurement at a certain temperature, frequency, and composition determines the value of ϑ in the binary mixture. The value of f_{AA} at a corresponding partial pressure of pure A is then substituted in (6.3) and f_{AB} is obtained. Since, roughly

$$f_{AB}=4N(1-m)\sigma_{AB}^2\xi_{AB}(\pi RT(M_A+M_B)/M_AM_B)^{\frac{1}{2}}e^{-A_{AB}/RT}, \tag{6.4}$$

the temperature variation of f_{AB} will permit a comparison of the steric and energetic effectiveness of collisions of the types AA and AB.

The next simplest binary mixture consists of two gases A and B each of which has a dispersive region in the frequency range under observation. If direct transmission of energy from the vibrations of A to those of B is not possible, the characteristic acoustical equation of this mixture will be given by (6.1) if $P=P_Am+P_B(1-m)$,

$$C_1=(C_{0A}-C_{\infty A})m, \quad C_2=(C_{0B}-C_{\infty B})(1-m),$$

$$\vartheta_1=\text{const.}_A/(f_{AA}+f_{AB})(1+e^{-x}\ )$$

and $\vartheta_1=\text{const.}_B/(f_{BB}+f_{AA})(1+e^{-x_B})$ are substituted. In view of what has been said in connection with (6.1) and (6.2) this requires no further discussion.

It should be pointed out, however, that this case provides the best means of investigating advantages due to vibrational resonance effects in collisions between unlike molecules. The probabilities f_{AA} and f_{BB} may be separately determined for the two pure components, and the analysis of measurements on such a mixture will therefore be at a great advantage over similar measurements in pure gases. By suitably varying the number of molecules in the excited states of the two molecular species with temperature, there appears to be a reasonable chance that significant information concerning vibrational resonance may be obtainable.

The behavior of a gas on illumination

If a gas is illuminated by light of wave-length suitable to raise it to a state of electronic excitation, the velocity of sound may change. The vibrational levels in the excited state will be, of course, quite different from those in the normal state. On radiating the gas may be expected to drop back to those vibrational levels in the normal state which disturb the relative separation of the nuclei as little as possible, in accord with the Franck-Condon principle. These will in general be of higher energy than those excited by collision. An alteration of the heat capacity and a change in the velocity of sound may therefore be expected. The situation so created is, of course, exceedingly complicated, and does not encourage general treatment. The number of molecules raised to the electronic state will also be only a very small fraction of

the total since the absorption coefficient for light is customarily small in gases. Nevertheless it may be possible to find a system which shows the effect of illumination to a marked extent with strong light sources and, in addition, possesses a sufficiently simple energy structure to permit its quantitative study without prohibitive labor. The possibility is mentioned here since experiments which are open to this interpretation have already been reported in the nondispersive region of several gases.[26] The study of a dispersive gas with a high optical absorption coefficient such as bromine might, in the dispersive region, yield valuable results if the optical conditions were quantitatively controlled.

[26] See for example, Kupper, Ann. d. Physik **43**, 905 (1912); Trautz, Zeits. f. Elektrochemie **18**, 513, 651 (1912).

* * * * * * * * * * * * * * * * *

It is hoped to make the next two parts of this series a detailed account of the experimental investigations on carbon dioxide, carbon disulphide, and binary mixtures of ethylene with other gases which have already been partially reported.[17]

It is a pleasure to acknowledge the assistance given throughout this study by Dr. Max Delbrück of the Kaiser Wilhelm-Institut für Radiochemie. Not only has he helped me to understand what Bourgin has written by correcting the immense number of printer's errors which occur in the work of this writer, but he has also frequently helped me to understand what I have written myself.

Thermodynamic Considerations on a New Kind of Transition Points

OTTO HALPERN, *Physics Department, New York University, University Heights*

(Received October 14, 1933)

According to previous experiments liquid helium shows a "transition point" with discontinuity in the values of its specific heat and expansion coefficient. It has not been possible, however, to observe the coexistence of the two "phases." This fact is here explained on a thermodynamic basis. Coexistence of two phases at a transition point can only be anticipated if the two phases exhibit a discontinuity in energy. An analogy is made with similar cases like that of the "transition point of superconductivity."

EXPERIMENTS[1] have recently led to the surprising result that liquid helium shows a "transition point" at extremely low temperatures. This transition point distinguishes itself from those with which we are usually acquainted by the fact that the two phases of helium exhibit no discontinuity as far as energy content or specific volume is concerned. They distinguish themselves from each other essentially by a jump in the values of the expansion coefficient $(\partial v/\partial T)_p$ and of the specific heat c_p. It has been found experimentally that this transition point also depends upon the pressure so that we should speak more strictly of a transition curve in the $p-T$ diagram. Furthermore, relations have been found to exist between the values for the discontinuity of the specific heat and the expansion coefficient. These empirical relations have been derived theoretically as a consequence of the second law by Keesom[1] and Ehrenfest.[2] We reproduce briefly the line of reasoning developed by Ehrenfest, since it will help to disclose the main theme of this short note more easily.

Let us denote by Z the thermodynamic potential of each helium phase; Z is a function of p and T. Ehrenfest considers the two modifications of helium as two different phases which when in equilibrium are characterized by the condition $Z_1=Z_2$. If we differentiate Z with respect to p and T, we obtain, $\partial Z/\partial p=v$ and $\partial Z/\partial T=-\rho$. Since there is no discontinuity in volume or energy (and therefore entropy) the first differential coefficients of the thermodynamic potential are continuous at the transition point. This does not hold true for the second derivatives.

$$\partial^2 Z/\partial T^2=-\partial/\partial T=-c_p/T,$$

$$\partial^2 Z/\partial T\partial p=\partial v/\partial T,$$

$$\partial^2 Z/\partial p^2=\partial v/\partial p.$$

Since as mentioned above the values for the entropies do not show a discontinuity at any point of the transition curve, we may take the total differential of both sides of the equation $S_1=S_2$ provided the differentials dp and dT are so chosen that we always remain on the transition curve. These special values of dp and dT shall be denoted by Dp and DT, following Ehrenfest.

We thus obtain

$$-(\partial v_1/\partial T)_p Dp+(c_{p1}/T)DT = -(\partial v_2/\partial T)_p Dp+(c_{p2}/T)DT$$

or

$$Dp/DT=(c_{p1}-c_{p2})/T(\partial v_1/\partial T-\partial v_2/\partial T),$$

a relation which has been established experimentally by Keesom. Several other similar relations can be derived which need not be discussed at present.

The question which has been raised by Ehrenfest, and which we shall try to answer is the following: Though we have treated liquid helium on both sides of the transition point as two different phases in a thermodynamical sense, and though the results derived from the equilibrium condition for the two phases are correct and in agreement with experiment, the observa-

[1] For references cf. W. H. Keesom, Proc. Amsterdam Acad. Sci., Sect. A **36**, 147 (1933).

[2] Ehrenfest, Proc. Amsterdam Acad. Sci., Sect. A **36**, 153 (1933).

tion of the coexistence of the two phases at the transition point as is usually possible at all other transition points has not been accomplished. What is the nature of the distinction between the helium transition point and the commonly observed transition points?

This question is readily answered if, for the sake of lucidity, we take as the independent variables of our thermodynamic system energy and volume instead of temperature and pressure. It then becomes obvious that a coexistence of two phases is only stable, i.e., observable, if these two phases show a *discontinuity in energy* (the *first* differential coefficient of the thermodynamic potential) and *not only in specific heat* (the *second* differential coefficient of the thermodynamic potential).

This will be clear from the following considerations. Let us plot, for example, the melting curve in the energy-volume diagram. Each point in this curve represents a certain equilibrium state with a definite ratio between the respective amounts of molten and solid substance present. To an infinitesimal variation of the energy there corresponds an infinitesimal displacement of the point on this curve or an infinitesimal change in this ratio. Temperature does not change unless we have reached the terminals of the curve at which points the whole substance is either solid or molten.

In the case of the helium transition point the situation is essentially different. Here all ratios of the composition of helium I and II correspond to *one* value of the energy. An infinitesimal change in energy would immediately eliminate one phase, and raise the temperature of the substance. This would hold true for every point in the transition curve, and not merely for the terminals, since there is no energy difference between the two phases. The coexistence of the two phases therefore never would be stable with respect to infinitesimal fluctuations in energy. This explains why the two phases could not be observed simultaneously.

It has already been pointed out by Ehrenfest that the discontinuity in the specific heat of helium shows much similarity to the discontinuity which is observable at the temperature at which a metal exhibits superconductivity. Here too it has not been possible to observe the coexistence of the two "metal phases." We can infer from the preceding line of reasoning that, at the transition point of superconductivity, no appreciable energy change occurs.

On Electrophoretic Mobilities and the Isoelectric Point of Protein-Coated Particles

OTTO HALPERN, *Physics Department, New York University, University Heights*

(Received October 14, 1933)

It is a fact known by experiment that quartz or glass particles of microscopic size suspended in a liquid and coated with a superficial layer of protein exhibit in almost every instance the same mobility in response to the application of an electric field. This indicates that their electrophoretic potential is independent of the size and shape of the particles. Furthermore, the isoelectric point of these particles coincides with the isoelectric point of the proteinic ions in the solution. Some very general thermodynamic arguments are given in this paper indicating that we should expect a similar behavior as a consequence of simple model properties.

THE difference between the electric potential as it enters into the formulae of thermodynamics, and the electrophoretic potential as determined by kinetic experiments (experiments of flow in general) is now clearly recognized.[1] While the first refers to the electrical work done in the transference of a charged particle from one phase to another, and therefore can be due to excess charges in one phase as well as to double layers rigidly connected with one phase, an electrokinetic potential, in the second place, can only be observed provided one of the phases, which are in motion relatively to each other show an excess charge. Keeping this in mind, it is not difficult to understand that in extreme cases the two potentials might change in an opposite sense, with the variation of the ionic concentration, as has been found experimentally by Freundlich,[2] and interpreted later by Stern.[3]

From the preceding remarks it should not be expected that it would be possible to draw any inferences relative to electrokinetic potentials from thermodynamics. This is generally true, although in the case at present under consideration it seems as if the combination of a simple assumption touching the model properties of the protein with general thermodynamic arguments permits of an understanding of experimentally discovered uniformities of behavior.

The theory of electrokinetic phenomena as developed by Helmholtz, Smoluchowski, and others shows that the electrokinetic potential ζ of a particle is connected with its mobility u under the influence of an electric field by the relation

$$\zeta = (C\eta/D)u \tag{1}$$

(η = coefficient of friction, D = dielectric constant). The numerical factor C is equal to 4π for particles the radius of which is large as compared to the thickness of the ionic atmosphere surrounding the particles.

Proteins dissolved in a solution of known pH act as amphoteric electrolytes; they are able to ionize by parting with either H^+ or OH^- ions. The dissociation equilibria are given by the relations

$$(C_+ C_{OH^-})/C = K_+, \tag{2}$$

$$(C_- C_{H^+})/C = K_-, \tag{3}$$

where the notation is self-explanatory.

Combining the relations (2) and (3) with the expression for the solubility product of water

$$C_{H^+} C_{OH^-} = K \tag{4}$$

we obtain as a condition for $C_+ = C_-$ the relation

$$C_{H^+} = [(K_-/K_+)K]^{\frac{1}{2}}. \tag{5}$$

The pH for which $C_+ = C_-$ is called the isoelectric point. At this isoelectric point the protein particles do not show any mobility in the electric field. This can be understood by a slight reinterpretation of the thermodynamic formulae. $C_+ = C_-$ does not mean that half of the

[1] For a treatment of the theory and experiments pertaining to electrokinetic phenomena, see H. A. Abramson, *Electrokinetic Phenomena*, Chem. Cat. Co., New York. (In Press.)

[2] H. Freundlich, *Kapillarchemie*, Leipzig, Vol. I, 1932.

[3] O. Stern, Zeits. f. Elektrochemie **30**, 508 (1924).

ions are positively, and the other half negatively charged, so that both types will move, only in opposite directions, but that every protein-particle is alternatingly positively and negatively charged, with the charge 0 in the time average. If the successive changes in the sign of the charge succeed each other fairly rapidly, the small motion in both directions becomes unobservable, the particles seeming to have the mobility 0.

If small particles of quartz or glass of microscopic dimensions but otherwise of arbitrary shape and size be immersed in the liquid it has been found by Abramson[4] that these particles coat themselves with a surface layer of protein, and subsequently show identical mobility in an electric field. Furthermore, the isoelectric point of the microscopic particles, i.e., the *pH* at which they show the mobility 0, coincides with the isoelectric point of molecularly dissolved protein. It is this general behavior which we are here trying to explain.

Let us denote by μ^i, $\mu_+{}^i$, $\mu_-{}^i$, respectively, the specific thermodynamic potential of the neutral, the positively, and the negatively ionized protein molecule in the surface of the particle denoted by i. We are considering for the following purposes the surface layer of adsorbed proteins as a phase in the thermodynamic sense, which assumption is very nearly correct, since the surface layer contains a large number of molecules. Removal or addition of a protein particle would influence the thermodynamic properties but very slightly. We shall furthermore assume that the *individual* specific thermodynamic potentials of the neutral, positively, and negatively charged protein particles apart from the concentration terms are all changed by the same amount through the process of adsorption. This assumption seems justified in view of the fact that the protein particles are giant molecules containing thousands of atoms, so that the change in free energy produced by adsorption will probably affect very little, if at all, the ionogenic properties of the protein. Our second assumption would be untenable if we were dealing, for example, with comparatively small molecules or ions, with which the adsorption would strongly influence the ionization properties.

[4] H. A. Abramson, J. Phys. Chem. **35**, 289 (1931); J. Gen. Physiol. **15**, 575 (1932).

In the case of thermodynamic equilibrium the following relation, as one out of the series of many similar, must hold

$$\mu_+{}^i + e\Delta_i = \mu_+ = \mu_+{}^k + e\Delta_k. \tag{6}$$

In Eq. (6), Δ stands for the electrical potential difference in the thermodynamic sense. Eq. (6) simply states that there must be equilibrium between any ion on the surface of any particle and the same ion in solution, and therefore between any two identical ions on the surfaces of two particles selected arbitrarily. According to our assumption as to the influence of adsorption we are justified in equalizing $\mu_+{}^i$ and $\mu_+{}^k$, so that Eq. (6) reduces to

$$\Delta_i = \Delta_k. \tag{7}$$

This relation was to be expected from the general behavior of "large electrodes" in electrochemistry.

Nevertheless, since Δ might still contain, as mentioned above, potential differences due to double layers rigidly connected to the particle, Eq. (7) *does not express* the equality of the electrophoretic potentials (i.e., the equality of the mobilities). To derive this second result we write the equilibrium condition Eq. (6) in the form

$$\mu_{+0}{}^i + RT \log C_+{}^i + e\Delta_i = \mu_{+0} + RT \log C_+, \tag{8}$$

$$\mu_{-0}{}^i + RT \log C_-{}^i - e\Delta_i = \mu_{-0} + RT \log C_-, \tag{9}$$

where we have assumed, as is certainly correct, that we are dealing with an ideal mixture. We subtract Eq. (9) from Eq. (8), and apply the resulting equation

$$\mu_{+0}{}^i - \mu_{-0}{}^i + RT \log (C_+{}^i/C_-{}^i) + 2e\Delta_i = \mu_{+0} - \mu_{-0} + RT \log (C_+/C_-) \tag{10}$$

to the isoelectric point

$$C_+ = C_-, \qquad C_+{}^i = C_-{}^i.$$

At this isoelectric point there is no more excess charge present on the particle (mobility O), so that any difference in potential would be due to double layers. Rewriting Eq. (10) we obtain

$$2e\Delta_i = \mu_{+0} - \mu_{+0}{}^i - (\mu_{-0} - \mu_{-0}{}^i), \tag{11}$$

in which the right-hand side is 0 (or at all

events very small), according to our second assumption. This shows for our model the absence of potential differences due to dipole layers, and therefore leads, with the help of (7), to the desired result: the potential difference due to excess charge and (therefore the mobility) is the same for all particles, independently of size and shape.

It becomes quite obvious in our presentation that the identity of the isoelectric point for free and adsorbed particles is by no means necessary, and in the case of proteins is due to the fact that adsorption apparently does not influence the ionogenic properties. In more general cases in which the dissociation is changed by adsorption the identity of the isoelectric point is not to be expected, and then our conclusions would become invalid.

The Thermal Dissociation of Cyanogen into Cyanide Radicals. A Correction

(J. Chem. Phys. 1, 432, 1933)

WE wish to correct a number of numerical errors which are present in our article of the above title.

On page 438, column 1, line 9, 567 should be written instead of 563, and on the following line 73 instead of 77. In the same column, lines 5 and 4 from the bottom, 234 should be corrected to 238; on the next line 22 changed to 27, and on the last line 153 to 152.

In the second column of the same page, on line 3, 147 must be changed to 146 and on line 32, 161 changed to 164.

The conclusions reached before remain substantially unaltered by these corrections.

G. B. KISTIAKOWSKY
H. GERSHINOWITZ
Harvard University

LETTERS TO THE EDITOR

This section will accept reports of new work, provided these are terse and contain few figures, and especially few halftone cuts. The Editorial Board will not hold itself responsible for opinions expressed by the correspondents. Contributions to this section must reach the office of the Managing Editor not later than the 15th of the month preceding that of the issue in which the letter is to appear. No proof will be sent to the authors. The usual publication charge ($3.00 per page) will not be made and no reprints will be furnished free.

The Formation of Formaldehyde by the Action of Ultraviolet Light on Carbon Dioxide and Water: An Application of the Allison Magneto-Optic Apparatus

Since the Allison[1] magneto-optic apparatus gives readings specific for each compound in solution, regardless of others that may be present, and since it will detect concentrations as low as 3 or 4 parts in 10^{12}, it should prove to be a powerful research tool in dealing with many problems in chemistry, physics, and biology. We have recently used this apparatus for the detection and quantitative determination of formaldehyde formed by the action of ultraviolet light on carbon dioxide and water. Formaldehyde gives minima at scale readings 21.83 and 21.92 Allison units. These minima are not found in pure water nor in carbonated water that has not been exposed to ultraviolet light. Upon placing the carbon dioxide saturated water in a quartz flask and irradiating the solution for several hours with the light from a mercury arc lamp, we obtained the light minima characteristic of formaldehyde. We have made ultraviolet light exposures up to 25 hours and find a uniform increase in the formaldehyde concentration with an increase in irradiation period. That this reaction is accompanied with a liberation of oxygen is indicated by the fact we have been able to oxidize (simultaneously with the formaldehyde formation) certain metallic ions from a lower to a higher valence. "Blanks" were run in all experiments to insure the absence of higher valence ions at the start.

This work is being continued and a full account will be published elsewhere upon completion of the investigation.

JOHN H. YOE
R. E. WINGARD

Cobb Chemical Laboratory,
University, Virginia
and
Department of Physics,
Alabama Polytechnic Institute,
Auburn, Alabama,
October 28, 1933.

[1] Allison and Murphy, J. Am. Chem. Soc. **52**, 3796 (1930); Allison, Ind. Eng. Chem., Anal. Edition **4**, 9 (1932); Allison, J. Chem. Ed. **10**, 71 (1933).

Helium Hydride Ion

The existence of helium hydride ion had been predicted on the basis of considerations regarding isoelectronic structures[1] and this ion has actually been found by the methods of mass-spectroscopy by Bainbridge.[2] Similar rare gas hydride ions have been found by Lukanow and Schütze.[3]

The structure of this ion has not yet been considered on the basis of wave mechanics and we thought the appropriate calculation should be made. The problem was considered in two ways:

(1) We supposed that the structure resulted from the interaction of an excited helium atom

$$\psi_1 = ae^{-\alpha r_1}; \qquad \psi_2 = (b+cr_2)e^{-\beta r_2}$$

with a proton. These ψ-functions for the two electrons of excited helium were found satisfactory when used to calculate the total energy of the helium atom ($2\,^3S$, 19.77 e.v.) with the values of the parameters α and β, 2 and 0.7, respectively. However, it was found that at all distances of separation of the atom and the proton repulsive forces predominated.

(2) We next assumed that the structure of helium hydride ion was made up of one helium ion and a hydrogen atom. Applying the method of Rosen[4] for the unsymmetrical system with $\alpha=2$, $\beta=1$ and $m=n=1$ which

[1] G. Glockler, D. L. Fuller and Charles P. Roe, J. Chem. Phys. **1**, 703 (1933).

[2] K. T. Bainbridge, Phys. Rev. **44**, 59 (1933).

[3] H. Lukanow and W. Schütze, Zeits. f. Physik **82**, 610 (1933).

[4] N. Rosen, Phys. Rev. **38**, 255 (1931).

implies that the ψ-functions used are purely hydrogenic for the normal state, we repeated the calculations. We evaluated all the integrals with exception of J_2 (page 272 of Rosen's paper) and now we find that there is an energy minimum of 8.1 e.v. at a distance of 1.3 Bohr hydrogen radii. The effect of the omitted integral J_2 will be to reduce the value of the energy.

The calculation of the integral J_2 mentioned above appears very laborious and at the present time it is not possible for us to spend the necessary time in its evaluation. However, we believe from physical considerations that its inclusion in the calculation will not change the main result of our work, i.e., that a minimum exists in the potential energy curve of helium hydride ion and that therefore wave mechanics agrees with experiment in making HeH^+ a stable structure.

It is a pleasure to thank Professor E. L. Hill of the Department of Physics for the advice given and the interest he has shown in these calculations.

GEORGE GLOCKLER
D. L. FULLER

University of Minnesota,
Minneapolis, Minnesota,
November 13, 1933.

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 1 NUMBER 12

DECEMBER, 1933

Published by the

AMERICAN INSTITUTE OF PHYSICS

Incorporated

LANCASTER, PA., AND NEW YORK, N. Y.

THE JOURNAL OF CHEMICAL PHYSICS

ADVISORY EDITORIAL BOARD

R. T. BIRGE
A. H. COMPTON
IRVING LANGMUIR
GILBERT N. LEWIS
A. A. NOYES
JOHN T. TATE

BOARD OF EDITORS

HAROLD C. UREY, *Managing Editor*

Associate Editors

DONALD H. ANDREWS
E. U. CONDON
FARRINGTON DANIELS
WHEELER P. DAVEY
DAVID M. DENNISON
F. G. DONNAN
R. H. FOWLER
KARL F. HERZFELD
C. N. HINSHELWOOD
T. R. HOGNESS
ERIC R. JETTE
FREDERICK G. KEYES
G. B. KISTIAKOWSKY
ELMER O. KRAEMER
C. H. KUNSMAN
VICTOR K. LA MER
WENDELL M. LATIMER
A. C. G. MITCHELL
W. ALBERT NOYES, JR.
LINUS PAULING
CHARLES P. SMYTH
HUGH S. TAYLOR

THE JOURNAL OF CHEMICAL PHYSICS is published monthly by the American Institute of Physics. Its purpose is to bridge a gap between journals of physics and journals of chemistry. The artificial boundary between physics and chemistry has now been in actual fact completely eliminated, and a large and active group is engaged in research which is as much the one as the other. It is to this group that the journal is expected to render its principal service and make its greatest appeal, both as a prompt and efficient medium for publication of research and as a convenient and carefully edited assemblage of pertinent information.

Manuscripts for publication should be submitted to Harold C. Urey, Havemeyer Hall, Columbia University, New York, New York.

Proof and all correspondence concerning papers in the process of publication should be addressed to the Editorial Secretary, American Institute of Physics, 11 East 38th Street, New York, New York. Subscriptions and orders for single copies should be addressed to Prince and Lemon Sts., Lancaster, Pa., or American Institute of Physics, 11 East 38th Street, New York, New York.

Subscription price $10.00 per year. Single copies $1.00.

The Journal of Chemical Physics is published monthly at Prince and Lemon Streets, Lancaster, Pennsylvania.

Entered as second-class matter December 30, 1932, at the Post Office at Lancaster, Pa., under the Act of August 24, 1912. Acceptance for mailing at special rate of postage, provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. & R.; authorized February 3, 1933.

THE AMERICAN INSTITUTE OF PHYSICS

INCORPORATED

GOVERNING BOARD

For The American Physical Society:

KARL T. COMPTON, *Chairman*
Massachusetts Institute of Technology

GEORGE B. PEGRAM, *Secretary*
Columbia University

JOHN T. TATE
University of Minnesota

For The Optical Society of America:

PAUL D. FOOTE
The Gulf Companies

LOYD A. JONES
Eastman Kodak Company

F. K. RICHTMYER
Cornell University

For The Acoustical Society of America:

HARVEY FLETCHER
Bell Telephone Laboratories

DAYTON C. MILLER
Case School of Applied Science

WALLACE WATERFALL
The Celotex Company

For The Society of Rheology:

E. C. BINGHAM
Lafayette College

WHEELER P. DAVEY
Pennsylvania State College

A. STUART HUNTER
du Pont Rayon Company

For The American Association of Physics Teachers:

HOMER L. DODGE
University of Oklahoma

PAUL E. KLOPSTEG
Central Scientific Company

FREDERIC PALMER
Haverford College

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*
JOHN T. TATE, *Adviser on Publications*
MADELINE M. MITCHELL, *Editorial Secretary*
Office: 11 East 38th Street, New York, New York

JOURNALS

THE PHYSICAL REVIEW
REVIEWS OF MODERN PHYSICS
PHYSICS
JOHN T. TATE, *Editor*
E. L. HILL, *Assistant Editor*
E. C. BINGHAM, *Rheology Editor*

THE JOURNAL OF THE OPTICAL SOCIETY OF AMERICA
THE REVIEW OF SCIENTIFIC INSTRUMENTS WITH PHYSICS NEWS AND VIEWS
F. K. RICHTMYER, *Editor*

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA
F. R. WATSON, *Editor*

THE JOURNAL OF CHEMICAL PHYSICS
HAROLD C. UREY, *Editor*

THE AMERICAN PHYSICS TEACHER
DUANE ROLLER, *Editor*

The American Institute of Physics is a membership corporation of which five societies are the founding members. These five societies are the principal national societies devoted to pure and applied physics in America. Their names, and those of their representatives who form the Governing Board of the Institute, are given above. The principal work of the Institute is the publication of journals sponsored by the societies or by itself. The names and the editors of the journals are also listed above.

The broad purpose of the Institute is to represent, in all matters of wide or common interest, the five thousand or more members and subscribers associated with the Founder Societies and the journals. It aims to advance the science and profession of physics, and to promote cooperation between pure research, the applied sciences and the industries. To achieve these aims, it is greatly in need of an endowment and is empowered by law to administer grants and funds of any kind.

CONTENTS

J. CHEM. PHYS. 1 (1933) DECEMBER

Derivation of Equations for Regular Solutions . . . . . J. H. HILDEBRAND AND S. E. WOOD 817

Technique for the Electrolytic Production of H^2H^2O . . . . . HUGH S. TAYLOR, HENRY EYRING AND ARTHUR A. FROST 823

Ground State of the Hydrogen Molecule . . . . . HUBERT M. JAMES AND ALBERT SPRAGUE COOLIDGE 825

Dielectric Constant Studies. I. An Improved Voltage Tuning Resonance Method and Its Application to Aqueous Potassium Chloride Solutions . . . . . J. GILBERT MALONE, A. L. FERGUSON AND L. O. CASE 836

Dielectric Constant Studies. II. The Drude Method Applied to Aqueous Solutions of Potassium Chloride . . . . . J. GILBERT MALONE, L. O. CASE AND A. L. FERGUSON 842

Heat of Formation of Binary Liquid Solutions from Their Liquid Components . . . . . J. J. COLEMAN AND FRANK E. E. GERMANN 847

Contact Potentials and the Effects of Unimolecular Films on Surface Potentials. I. Films of Acids and Alcohols . . . . . WILLIAM D. HARKINS AND E. K. FISCHER 852

Acoustical Studies. II. The Behavior of a Gas with Several Independent Internal Energy States . . . . . WILLIAM T. RICHARDS 863

Thermodynamic Considerations on a New Kind of Transition Points . . . . . OTTO HALPERN 880

On Electrophoretic Mobilities and the Isoelectric Point of Protein-Coated Particles . . . . . OTTO HALPERN 882

Thermal Dissociation of Cyanogen into Cyanide Radicals. A Correction . . . . . G. B. KISTIAKOWSKY AND H. GERSHINOWITZ 885

Letters to the Editor:

Formation of Formaldehyde by the Action of Ultraviolet Light on Carbon Dioxide and Water: An Application of the Allison Magneto-Optic Apparatus . . . . . JOHN H. YOE AND R. E. WINGARD 886

Helium Hydride Ion . . . . . GEORGE GLOCKLER AND D. L. FULLER 886

Author Index to Volume 1 888

Analytic Subject Index to Volume 1 892

LANCASTER PRESS, INC., LANCASTER, PA.